BLEJSKE DELAVNICE IZ FIZIKE LETNIK 4, ŠT. 2–3
BLED WORKSHOPS IN PHYSICS VOL. 4, NO. 2–3

ISSN 1580–4992

Proceedings to the Euroconference on Symmetries Beyond the Standard Model

Portorož, July 12 – 17, 2003

(Part 2 of 2)

Edited by

Norma Mankoč Borštnik[1,2]
Holger Bech Nielsen[3]
Colin D. Froggatt[4]
Dragan Lukman[2]

[1]*University of Ljubljana,* [2]*PINT,* [3] *Niels Bohr Institute,* [4] *Glasgow University*

DMFA – ZALOŽNIŠTVO
LJUBLJANA, DECEMBER 2003

The *Euroconference on Symmetries Beyond the Standard Model*, 12.–17. July 2003, Portorož, Slovenia

was organized by

European Science Foundation, EURESCO office, Strassbourg

and sponsored by

Ministry of Education, Science and Sport of Slovenia

Luka Koper d.d., Koper, Slovenia

Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana

Primorska Institute of Natural Sciences and Technology, Koper

Scientific Organizing Committee

Norma Mankoč Borštnik (Chairperson), Ljubljana and Koper, Slovenia

Holger Bech Nielsen (Vice Chairperson), Copenhagen, Denmark

Colin D. Froggatt, Glasgow, United Kingdom

Loriano Bonora, Trieste, Italy

Roman Jackiw, Cambridge, Massachussetts, USA

Kumar S. Narain, Trieste, Italy

Julius Wess, Munich, Germany

Donald L. Bennett, Copenhagen, Denmark

Pavle Saksida, Ljubljana, Slovenia

Astri Kleppe, Oslo, Norway

Dragan Lukman, Koper, Slovenia

Contents

Bled Workshops
in Physics
Vol. 4, No. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 138)*
Portorož, Slovenia, July 12-17, 2003

Supersymmetric Grandunification and Fermion Masses

Borut Bajc

Jožef Stefan Institute, Jamova 39, 1001 Ljubljana, Slovenia
borut.bajc@ijs.si

Abstract. A short review of the status of supersymmetric grand unified theories and their relation to the issue of fermion masses and mixings is given.

1 Why Grandunification?

There are essentially three reasons for trying to build grand unified theories (GUTs) beyond the standard model (SM).

- why should strong, weak and electromagnetic couplings in the SM be so different despite all corresponding to gauge symmetries?
- there are many disconnected matter representations in the SM (3 families of L, e^c, Q, u^c, d^c)
- quantization of electric charge (in the SM model there are two possible explanations - anomaly cancellation and existence of magnetic monopoles - both are naturally embodied in GUTs)

2 How to check a GUT?

I will present here a very short review of some generic features, predictions and drawbacks of GUTs. Details of some topics will be given in the next section.

2.1 Gauge coupling unification

This is of course a necessary condition for any GUT to work. As is well known, the SM field content plus the desert assumption do not lead to the unification of the three gauge couplings. However, the idea of low-energy supersymmetry (susy), i.e. the minimal supersymmetric standard model (MSSM) instead of the SM at around TeV and again the assumption of the desert gives a quite precise unification of gauge couplings at $M_{GUT} \approx 10^{16}$ GeV [1]. Clearly there is no a-priori reason for three functions to cross in one point, so this fact is a strong argument for supersymmetry. One gets two bonuses for free in this case. First, the hierarchy problem gets stabilized, although not really solved, since the famous doublet-triplet (DT) problem still remains. Secondly, at least in principle one can

get an insight into the reasons for the electroweak symmetry breaking: why the Higgs (other bosons in MSSM) mass squared is negative (positive) at low energy [2].

2.2 Fermion masses and mixings

Although GUTs are not theories of flavour, they bring constraints on the possible Yukawas. In the MSSM the Yukawa sector is given by

$$W_Y = HQ^T Y_U u^c + \overline{H} Q^T Y_D d^c + \overline{H} L^T Y_E e^c \ , \tag{1}$$

and the complex 3×3 generation matrices $Y_{U,D,E}$ are arbitrary. However, in a GUT the matter fields Q, L, u^c, d^c, e^c fields live together in bigger representations, so one expects relations between quark and lepton Yukawa matrices.

Take for example the SO(10) GUT. All the MSSM matter fields of each generation live in the same representation, the 16 dimensional spinor representation, which contains and thus predicts also the right-handed neutrino. At the same time the minimal Higgs representation, the 10 dimensional representation contains both doublets H and $\overline{H}$ of the MSSM (plus one color triplet and one antitriplet). The only renormalizable SO(10) invariant one can write down is thus

$$W_Y = 10_H 16 Y_{10} 16 \ , \tag{2}$$

which is however too restrictive, since it gives on top of the well satisfied (for large $\tan\beta$) relation $y_b = y_\tau = y_t$ for the third generation, the much worse predictions for the first two generations ($y_s = y_\mu = y_c$ and $y_d = y_e = y_u$) and no mixing ($\theta_c = 0$) at all.

How to improve the fit? Let us mention two possibilities:

(1) Introduce new Higgs representations: although another 10_H can help with the mixing, the experimentally wrong relations $m_d = m_e$ and $m_s = m_\mu$ still occur, because the two bi-doublets in the two 10_H leave invariant the Pati-Salam SO(6)=SU(4)$_C$, so the leptons and quarks are still treated on the same footing. So the idea to pursue is to introduce bidoublets which transform nontrivially under the Pati-Salam SU(4) color. This can be done for example by introducing a $\overline{126}_H$, which couples to matter as $\Delta W_Y = \overline{126}_H 16 Y_{126} 16$ and which gets a nontrivial vev in the $(2,2,15)_H$ SU(3) color singlet direction [3,4].

(2) Another possibility is to include the effects of nonrenormalizable operators. These operators can cure the problem and at the same time ease the proton decay constraints. The drawback is the loss of predictivity.

2.3 Proton decay

This issue is connected to

(1) R-parity. It is needed to avoid fast proton decay. At the nonrenormalizable level one could for example have terms leading to R-parity violation of the type $16^3 16_H / M_{Pl}$. For this reason it is preferable to use the 126_H representation instead of the 16_H. It is possible to show that such a SO(10) with 126_H has an exact R-parity [5] at all energies without the introduction of further symmetries.

(2) DT splitting problem: Higgs $SU(2)_L$ doublets and $SU(3)_C$ triplets live usually in the same GUT multiplet; but while the $SU(2)_L$ doublets are light ($\approx M_W \ll M_{GUT}$), the $SU(3)_C$ triplets should be very heavy ($\geq M_{GUT}$) to avoid a too fast proton decay. For example, the proton lifetime in susy is proportional to M_T^2, which can give a lower limit to the triplet mass [6], although this limit depends on the yet unknown supersymmetry breaking sector [7].

The solutions to the DT problem depend on the gauge group considered, but in general models that solve it are not minimal and necessitate of additional Higgs sectors. For example the missing partner mechanism [8] in SU(5) needs at least additional 75_H, 50_H and $\overline{50}_H$ representations. The same is true for the missing vev mechanism in SO(10) [9], where the 45_H and extra 10_H Higgses must be introduced. Also the nice idea of GIFT (Goldstones Instead of Fine Tuning) [10] can be implemented only by complicated models, while discrete symmetries for this purpose can be used with success only in connection with non-simple gauge groups [11]. Of course, although not very natural, any GUT can "solve" the problem phenomenologically, i.e. simply fine-tuning the model parameters.

Clearly, whatever is the solution to the DT problem, the proton lifetime depends in a crucial (powerlike) way on the triplet mass. And this mass can be difficult to determine from the gauge coupling unification condition even in specific models because of the unknown model parameters [12] or use of high representations [13]. On top of this there can be large uncertainties in the triplet Yukawa couplings [14]. All this, together with the phenomenologically completely unknown soft susy breaking sector, makes unfortunately proton decay not a very neat probe of supersymmetric grandunification [7]. Of course, if for some reason the DT mechanism is so efficient to make the $d = 6$ operators dominant (for a recent analysis in some string-inspired models see [15]), then the situation could be simpler to analyse [16], although many uncertainties due to fermion mixing matrices still exist in realistic nonminimal models [17]. Unfortunately there is little hope to detect proton decay in this case, unless M_{GUT} is lower than usual [18].

2.4 Magnetic monopoles

Since magnetic monopoles are too heavy to be produced in colliders, the only hope is to find them as relics from the cosmological GUT phase transitions. Their density however strongly depends on the cosmological model considered. Unfortunately, the Rubakov-Callan effect [19] leads to the non observability of GUT monopoles, at least in any foreseeable future. Namely, these monopoles are captured by neutron stars and the resulting astrophysical analyses [20] limits the monopole flux at earth twelve orders of magnitude below the MACRO limit [21].

This is very different from the situation in the Pati-Salam (PS) theory. Even in the minimal version the PS scale can be much lower than the GUT scale [22], as low as 10^{10} GeV. the resulting monopoles are then too light to be captured by neutron stars and their flux is not limited due to the Rubakov-Callan effect. Furthermore, MACRO results are not applicable for such light monopoles [21].

2.5 Low energy tests

There are many different possible tests at low-energy, like for example the flavour changing neutral currents (see for example [23]) or the electric dipole moments [24]. In the latter case the exact value of the triplet mass is much less important than in proton decay, but the uncertainties due to the susy breaking sector are still present. In some of these tests like neutron-antineutron oscillation we can get positive signatures only for specific models due to very high dimensional operators involved [25].

3 Fermion masses and mixings

The regular pattern of 3 generations suggests some sort of flavour symmetry.

One way, and the most ambitious one, is to consider the flavour symmetry group as part (subgroup) of the grand unified gauge symmetry (described by a simple group). In such an approach all three generations come from the same GUT multiplet. For example, in SU(8) the 216 dimensional representation gets decomposed under its SU(5) subgroup into three copies (generations) of $\overline{5}$ and 10 with additional SU(5) multiplets. Similarly, in the SO(18) GUT, the 256 dimensional spinorial representation is nothing else than 8 generations of $(16+\overline{16})$ in the SO(10) language. The problem in all these theories is what to do with all the extra light particles [26].

Another possibility is to consider the product of the flavour (or, in general, extra) symmetry with the GUT symmetry (simple) group. In the context of SO(10) GUTs most of them use small representations for the Higgses, like 16_H, $\overline{16}_H$ and 45_H. The philosphy is to consider all terms allowed by symmetry, also nonrenormalizable. The DT problem can be naturally solved by some version of the missing vev mechanism, which however means that many multiplets are usually needed. Such models are quite successfull [27], although the assumed symmetries are a little bit ad-hoc. There is also a huge number of different models with almost arbitrary flavour symmetry group, but unfortunately there is no room to describe them here (see for example the recent review [28]).

What we will consider in the following is instead a SO(10) GUT with no extra symmetry at all. We want to see how far we can go with just the grand unified gauge symmetry alone. To ensure automatic R-parity, we are forced not to use the 16_H and $\overline{16}_H$ Higgses, but instead a pair of 126_H and $\overline{126}_H$ (5 index antisymmetric representations, one self-dual, the other anti-self-dual; both of them are needed in order not to break susy at a large scale). In fact under R-parity the bosons of 16 are odd, while those of 126 are even, since

$$R=(-1)^{3(B-L)+2S} \tag{3}$$

[29], and the relevant vev in the SU(5) singlet directions have $B-L=1$ for 16_H (ν^c), while it has $B-L=2$ for 126 (the mass of ν^c).

So the rules of the game are: stick to renormalizable operators only, consider SO(10) as the only symmetry of the model, take the minimal number of multiplets (it does not mean the minimal number of fields!) that is able to give the correct

symmetry breaking pattern SO(10)→SU(3)×SU(2)×U(1). Such a theory is given by [30] (see however [31] for a similar approach): on top of the usual three generations of 16 dimensional matter fields, it contains four Higgs representations: 10_H, 126_H, $\overline{126}_H$ and 210_H (4 index antisymmetric). It has been shown recently [32] that this theory is also the minimal GUT, i.e. it has the minimal number of model parameters, being still perfectly realistic (not in contradiction with any experiment).

As we have seen, the $\overline{126}_H$ multiplet is needed both to help the 10_H multiplet in fitting the fermion masses and mixings, and for giving the mass to the right-handed neutrino without explicitly breaking R-parity. Let us now see, why the 210_H representation is needed.

The Yukawa sector is given by

$$W_Y = 10_H 16 Y_{10} 16 + \overline{126}_H 16 Y_{126} 16 \; . \tag{4}$$

The fields decompose under the $SU(2)_L \times SU(2)_R \times SU(4)_C$ subgroup as

$$10_H = (2,2,1) + (1,1,6) \; , \tag{5}$$
$$16 = (2,1,4) + (1,2,\overline{4}) \; , \tag{6}$$
$$\overline{126}_H = (1,3,10) + (3,1,\overline{10}) + (2,2,15) + (1,1,6) \; . \tag{7}$$

The right-handed neutrino ν^c lives in $(1,2,\overline{4})$ of 16, so it can get a large mass only through the second term in (4):

$$M_{\nu_R} = \langle (1,3,10)_{\overline{126}} \rangle Y_{126} \; , \tag{8}$$

where $\langle (1,3,10) \rangle$ is the scale of the $SU(2)_R$ symmetry breaking M_R, which we assume to be large, $\mathcal{O}(M_{GUT})$.

In order to get realistic masses we need

$$\langle (2,2,1)_{10} \rangle = \begin{pmatrix} v_{10}^d & 0 \\ 0 & v_{10}^u \end{pmatrix} \neq 0 \; , \tag{9}$$
$$\langle (2,2,15)_{\overline{126}} \rangle = \begin{pmatrix} v_{126}^d & 0 \\ 0 & v_{126}^u \end{pmatrix} \neq 0 \; , \tag{10}$$

which contribute to the light fermion masses as

$$M_U = v_{10}^u Y_{10} + v_{126}^u Y_{126} \; , \tag{11}$$
$$M_D = v_{10}^d Y_{10} + v_{126}^d Y_{126} \; , \tag{12}$$
$$M_{\nu_D} = v_{10}^u Y_{10} - 3 v_{126}^u Y_{126} \; , \tag{13}$$
$$M_E = v_{10}^d Y_{10} - 3 v_{126}^d Y_{126} \; . \tag{14}$$

The factor of -3 for leptons in the contribution from $\overline{126}_H$ comes automatically from the fact that the $SU(3)_C$ singlet in the adjoint 15 of $SU(4)_C$ is in the

$B-L$ direction diag$(1,1,1,-3)$. This is clearly absent in the contribution from 10_H, which is a singlet under the full $SU(4)_C$.

The light neutrino mass comes through the famous see-saw mechanism [33]. From

$$W = \frac{1}{2}\nu^{cT} M_{\nu_R} \nu^c + \nu^{cT} M_{\nu_D} \nu_L + ... \tag{15}$$

one can integrate out the heavy right-handed neutrino ν^c obtaining the effective mass term for the light neutrino states $M_N = -M_{\nu_D}^T M_{\nu_R}^{-1} M_{\nu_D}$. As we will now see, there is another contribution in our minimal model.

(1) We saw that both $\langle (2,2,1)_{10} \rangle$ and $\langle (2,2,15)_{126} \rangle$ need to be nonzero and obviously $\mathcal{O}(M_W)$. With 10_H, 126_H and $\overline{126}_H$ Higgses one can write only two renormalizable invariants:

$$W_H = \frac{1}{2} M_{10} 10_H^2 + M_{126} 126_H \overline{126}_H \ , \tag{16}$$

where $M_{10}, M_{126} \approx \mathcal{O}(M_{GUT})$ or larger due to proton decay constraints. So the mass term looks like

$$\frac{1}{2}(10_H, 126_H, \overline{126}_H) \begin{pmatrix} M_{10} & 0 & 0 \\ 0 & 0 & M_{126} \\ 0 & M_{126} & 0 \end{pmatrix} \begin{pmatrix} 10_H \\ 126_H \\ \overline{126}_H \end{pmatrix} \ . \tag{17}$$

Clearly all the doublets have a large positive mass, so their vev must be zero. Even fine-tuning cannot solve the DT problem in this case. So the idea to overcome this obstacle is to mix in some way 10_H with $\overline{126}_H$ (126_H), and after that fine-tune to zero one combination of doublet masses. So the new mass matrix should look something like

$$\begin{pmatrix} M_{10} & x & y \\ x & 0 & M_{126} \\ y & M_{126} & 0 \end{pmatrix} \tag{18}$$

with x, y denoting such mixing. The light Higgs doublets will thus be linear combinations of the fields in $(2,2,1)_{10}$ and $(2,2,15)_{126_H, \overline{126}_H}$ and this will get a nonzero vev after including the soft susy breaking masses.

(2) The minimal representation that can mix 10 and $\overline{126}$ is 210, as can be seen from $10 \times \overline{126} = 210 + 1050$. 210 is a 4 index antisymmetric SO(10) representation, which decomposes under the Pati-Salam subgroup as

$$210 = (1,1,1)+(1,1,15)+(1,3,15)+(3,1,15)+(2,2,6)+(2,2,10)+(2,2,\overline{10}) \ . \tag{19}$$

Of course one can now add other renormalizable terms to (16), and all such new terms are (in a symbolic notation)

$$\Delta W_H = 210_H^3 + 210_H^2 + 210_H 126_H \overline{126}_H + 210_H 10_H 126_H + 210_H 10_H \overline{126}_H \ . \tag{20}$$

The last two terms are exactly the ones needed for the mixings between 10_H and 126_H ($\overline{126}_H$), i.e. contributions to x, y in (18). It is possible to show that $W_H + \Delta W_H$ are just enough for SO(10)→SM. In the case of single-step breaking one thus has

$$\begin{aligned} \langle (1,1,1)_{210} \rangle \approx \langle (1,1,15)_{210} \rangle \approx \langle (1,3,15)_{210} \rangle \approx \\ \langle (1,3,\overline{10})_{126} \rangle \approx \langle (1,3,10)_{\overline{126}} \rangle \approx M_{GUT} \ . \end{aligned} \tag{21}$$

(3) Now however there are five bidoublets that mix, since $(2,2,10)$ and $(2,2,\overline{10})$ from 210_H also contribute. To be honest, there is only one neutral component in each of these last two bidoublets, since their $B-L$ equals ± 2, so the final mass matrix for the Higgs doublets is 4×4. Only one eigenvalue of this matrix needs to be zero, and this can be achieved by fine-tuning. Each of the two Higgs doublets of the MSSM is thus a linear combination of 4 doublets, each of which has in general a vev of order $\mathcal{O}(M_W)$:

$$\begin{aligned} \langle (2,2,1)_{10} \rangle \approx \langle (2,2,15)_{\overline{126}} \rangle \approx M_W \approx \\ \langle (2,2,15)_{126} \rangle \approx \langle (2,2,10)_{210} \rangle \approx \langle (2,2,\overline{10})_{210} \rangle \ . \end{aligned} \tag{22}$$

This mixing is nothing else than the susy version of [3,4].

(4) Due to all these bidoublet vevs, a $SU(2)_L$ triplet will also get a tiny but nonzero vev. Applying the susy constraint $F_{(3,1,10)_{126}} = 0$ to

$$W = M_{126}(3,1,10)_{126}(3,1,\overline{10})_{\overline{126}} + (2,2,1)_{10}(2,2,\overline{10})_{210}(3,1,10)_{126} + ... \tag{23}$$

one immediately gets

$$\langle (3,1,\overline{10})_{\overline{126}} \approx \frac{\langle (2,2,1)_{10} \rangle \langle (2,2,\overline{10})_{210} \rangle}{M_{126}} \approx \frac{M_W^2}{M_{GUT}} \neq 0 \ . \tag{24}$$

This effect is just the susy version of [34].

(5) Since ν lives in $(2,1,4)_{16}$, the second term in (4) gives among others also a term $(3,1,\overline{10})_{\overline{126}}(2,1,4)_{16} Y_{126} (2,1,4)_{16}$, which contributes to the light neutrino mass. So all together one gets for the light neutrino mass (c is a model dependent dimensionless parameter)

$$M_N = -M_{\nu_D}^T M_{\nu_R}^{-1} M_{\nu_D} + c \frac{M_W^2}{M_{GUT}} Y_{126} \ . \tag{25}$$

The first term is called the type I (or canonical) see-saw and is mediated by the $SU(2)_L$ singlet ν^c, while the second is the type II (or non-canonical) see-saw, and is mediated by the $SU(2)_L$ triplet.

Equations (11), (12), (13), (14), (8) and (25) are all we need in the fit of known fermion masses and mixings and predictions of the unknown ones. A possible procedure is first to trade the matrices Y_{10} and Y_{126} for M_U and M_D. The remaining freedom in M_U and M_D is still enough to fit M_E. But then some predictions

in the neutrino sector are possible. For this sector we need to reproduce the experimental results $(\theta_l)_{12,23} \gg (\theta_q)_{12,23}$ and $(\theta_l)_{13}$ small. The degree of predictivity of the model however depends on the assumptions regarding the see-saw and on the CP phases.

The first approach was to consider models in which type I dominates. It was shown that such models predict a small atmospheric neutrino mixing angle $\theta_{atm} = (\theta_l)_{23}$ if the CP phases are assumend to be small [4,35]. On the other hand, a large atmospheric neutrino mixing angle can be also large, if one allows for arbitrary CP phases and fine-tune them appropriately [36].

A completely different picture emerges if one assumes that type II see-saw dominates. In this case even without CP violation one can naturally have a large atmospheric neutrino mixing angle, as has been first emphasized for the approximate case of second and third generations only in [37,38]. In the three generation case the same result has been confirmed [39]. On top of this, a large solar neutrino mixing angle and a prediction of $U_{e3} \approx 0.15 \pm 0.01$ (close to the upper experimental limit) have been obtained [39]. Even allowing for general CP violation does not invalidate the above results: although the error bars are larger, the general picture of large atmospheric and solar neutrino mixing angles and small U_{e3} still remains valid [40].

It is possible to understand why type II see-saw gives so naturally a large atmospheric mixing angle. In type II the light neutrino mass matrix (25) is proportional to Y_{126}. From (12) and (14) one can easily find out, that $Y_{126} \propto M_D - M_E$, from which one gets [41]

$$M_N \propto M_D - M_E \ . \tag{26}$$

As a warm-up let us take the approximations of just (a) two generations, the second and the third, (b) neglect the masses of the second generation with respect of the third ($m_{s,\mu} \ll m_{b,\tau}$) and (c) assume that M_D and M_E has small mixings (this amounts to say, that in the basis of diagonal charged lepton mass, the smallness of the $(\theta_q)_{23} = \theta_{cb}$ is not caused by accidental cancellation of two large numbers). In this approximate set-up one gets

$$M_N \propto \begin{pmatrix} 0 & 0 \\ 0 & m_b - m_\tau \end{pmatrix} \ . \tag{27}$$

This is, in type II see-saw there is a correspondence between the large atmospheric angle and $b - \tau$ unification [38].

Remember here that $b - \tau$ Yukawa unification is no more automatic, since we have also $\overline{126}_H$ Higgs on top of the usual 10_H. It is however quite well satisfied phenomenologically.

One can do better: still take $m_{s,\mu} \approx 0$, but allow a nonzero quark mixing. In this case the atmospheric mixing angle is

$$\tan 2\theta_{atm} = \frac{\sin 2\theta_{cb}}{2 \sin^2 \theta_{cb} - \frac{m_b - m_\tau}{m_b}} \ . \tag{28}$$

Since $\theta_{cb} \approx \mathcal{O}(10^{-2})$, one again finds out the correlation between the large atmospheric mixing angle and $b - \tau$ unification at the GUT scale.

The result can be confirmed of course also for finite $m_{s,\mu,c}$, although not in a so simple and elegant way.

Of course there are many other models that predict and/or explain a large atmospheric mixing angle (for a recent review see for example [42]). What is surprising here is, however, that no other symmetry except the gauge SO(10) is needed whatsoever.

4 The minimal model

As we saw in the previos section, one can correctly fit the known masses and mixings, get some understanding of the light neutrino mass matrix, and obtain some new predictions. What we would like to show here is that the model considered above has less number of model parameters than any other GUT, and can be then called the minimal realistic supersymmetric grand unified theory (even more minimal than SU(5)!) [32].

The Higgs sector described by (16) and (20) contains 10 real parameters (7 complex parameters minus four phase redefinitions due to the four complex Higgs multiplets involved). The Yukawa sector (4) has two complex symmetric matrices, one of which can be always made diagonal and real by a unitary transformation of 16 in generation space. So what remains are 15 real parameters. There is on top of this also the gauge coupling, so all together 26 real parameters in the supersymmetric sector of renormalizable SO(10) GUT with three copies of matter 16 and Higgses in the representations 10_H, 126_H, $\overline{126}_H$ and 210_H. We will not consider the susy breaking sector, since this is present in all supersymmetric theories, GUTs or not.

Before comparing with other GUTs, for example SU(5), let us count the number of model parameters in MSSM. There are 6 quark masses, 3 quark mixing angles, 1 quark CP phase, 6 lepton masses, 3 lepton mixing angles and 3 lepton CP phases (assuming Majorana neutrinos). On top of this, there are 3 gauge couplings and the real μ parameter. Thus, all together, again 26 real parameters. They are however distributed differently, so that in the Yukawa sector there are only 15 parameters in our minimal SO(10) GUT, which has to fit 22 MSSM (at least in principle) measurable low-energy parameters. Although in this fitting also few vevs that contain parameters from the Higgs and susy breaking sector play a role, the minimal SO(10) is nevertheless predictive.

One can play with other SO(10) models: the renormalizable ones need more representations and thus have more invariants, while the nonrenormalizable ones (those that use 16_H instead of 126_H) have a huge number of invariants, some of which must be very small due to R-symmetry constraints. Of course, with some extra discrete, global or local symmetry, one can forbid these unpleasant and dangerous terms, remaining even with a small number of parameters, but as we said, this is not allowed in our scheme, in which we want to obtain as much information as possible just from GUT gauge symmetry (and renormalizability).

The simplicity of the minimal renormalizable supersymmetric SU(5) looks as if the number of parameters here could be smaller than in our previous example. What however gives a large number of parameters is the fact, that SU(5) is not

particularly suitable for the neutrino sector. In fact, one can play and find out, that the minimal SU(5) with nonzero neutrino masses is obtained adding the two index symmetric 15_H and $\overline{15}_H$, and the number of model parameters comes out to be 39, i.e. much more than in the minimal SO(10).

5 Conclusion

Before talking about flavour symmetries it is important first to know, what we can learn from just pure GUTs. The minimal GUT is a SO(10) gauge theory with representations 10_H, 126_H, $\overline{126}_H$, 210_H and three generations of 16. Such a realistic theory is renormalizable and no extra symmetries are needed. It can fit the fermion masses and mixings, and can give an interesting relation between $b - \tau$ Yukawa unification and large atmospheric mixing angle. It has a testable prediction for U_{e3}. Due to the large representations involved, it is not asymptotically free, which means that it predicts some new physics below M_{Pl}.

There are many virtues of this minimal GUT. As in any SO(10) all fermions of one generation are in the same representation and the right-handed neutrino is included automatically, thus explaining the tiny neutrino masses by the see-saw mechanism. Employing 126_H instead of 16_H to break $B - L$ maintains R-symmetry exact at all energies. It is economical, it employs the minimal number of multiplets and parameters, and thus it is maximally predictive. It gives a good fit to available data and gives a framework to better understand the differences between the mixings in the quark and lepton sectors.

There are of course also some drawbacks. First, in order to maintain predictivity, one must believe in the principle of renormalizability, although the suppressing parameter in the expansion M_{GUT}/M_{Planck} is not that small. Of course, in supersymmetry these terms can be small and stable, but this choice is not natural in the 't Hooft sense. Second, the DT splitting problem is here, and attempts to solve it require more fields [43]. Finally, usually it is said that 126 dimensional representations are not easy to get from superstring theories, although we are probably far from a no-go theorem.

There are many open questions to study in the context of the minimal SO(10), let me mention just few of them. First, proton decay: although it is generically dangerous, it is probably still possible to fit the data with some fine-tuning of the model parameters as well as of soft susy breaking terms. An interesting question is whether the model is capable of telling us which type of see-saw dominates. If it is type I or mixed, can it still give some testable prediction for U_{e3}? Also, gauge coupling unification should be tested in some way, although large threshold corrections could be nasty [13]. And finally, is there some hope to solve in this context or minimal (but still predictive) extensions the doublet-triplet splitting problem?

Acknowledgements

It is a pleasure to thank the organizers for the well organized and stimulating conference. I am grateful to Charan Aulakh, Pavel Fileviez Perez, Alejandra Melfo,

Goran Senjanović and Francesco Vissani for a fruitful collaboration. I thank Goran Senjanović also for carefully reading the manuscript and giving several useful advises and suggestions. This work has been supported by the Ministry of Education, Science, and Sport of the Republic of Slovenia.

References

1. S. Dimopoulos, S. Raby and F. Wilczek, Phys. Rev. D **24** (1981) 1681; L. E. Ibañez and G. G. Ross, Phys. Lett. B **105** (1981) 439; M. B. Einhorn and D. R. Jones, Nucl. Phys. B **196** (1982) 475; W. J. Marciano and G. Senjanović, Phys. Rev. D **25** (1982) 3092.
2. K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita, Prog. Theor. Phys. **68** (1982) 927 [Erratum-ibid. **70** (1983) 330]; L. Alvarez-Gaume, J. Polchinski and M. B. Wise, Nucl. Phys. B **221** (1983) 495.
3. G. Lazarides, Q. Shafi and C. Wetterich, Nucl. Phys. B **181** (1981) 287.
4. K. S. Babu and R. N. Mohapatra, Phys. Rev. Lett. **70** (1993) 2845 [arXiv:hep-ph/9209215].
5. C. S. Aulakh, K. Benakli and G. Senjanović, Phys. Rev. Lett. **79** (1997) 2188 [arXiv:hep-ph/9703434]; C. S. Aulakh, A. Melfo and G. Senjanović, Phys. Rev. D **57** (1998) 4174 [arXiv:hep-ph/9707256]; C. S. Aulakh, A. Melfo, A. Rašin and G. Senjanović, Phys. Lett. B **459** (1999) 557 [arXiv:hep-ph/9902409].
6. J. Hisano, H. Murayama and T. Yanagida, Nucl. Phys. B **402** (1993) 46 [arXiv:hep-ph/9207279]; T. Goto and T. Nihei, Phys. Rev. D **59** (1999) 115009 [arXiv:hep-ph/9808255]; H. Murayama and A. Pierce, Phys. Rev. D **65** (2002) 055009 [arXiv:hep-ph/0108104].
7. B. Bajc, P. F. Perez and G. Senjanović, Phys. Rev. D **66** (2002) 075005 [arXiv:hep-ph/0204311] and arXiv:hep-ph/0210374.
8. A. Masiero, D. V. Nanopoulos, K. Tamvakis and T. Yanagida, Phys. Lett. B **115** (1982) 380.
9. S. Dimopoulos and F. Wilczek, Print-81-0600 (SANTA BARBARA); K. S. Babu and S. M. Barr, Phys. Rev. D **48** (1993) 5354 [arXiv:hep-ph/9306242] and Phys. Rev. D **65** (2002) 095009 [arXiv:hep-ph/0201130].
10. see for example Z. Berezhiani, C. Csaki and L. Randall, Nucl. Phys. B **444** (1995) 61 [arXiv:hep-ph/9501336], and references therein.
11. R. Barbieri, G. R. Dvali and A. Strumia, Phys. Lett. B **333** (1994) 79 [arXiv:hep-ph/9404278]; S. M. Barr, Phys. Rev. D **55** (1997) 6775 [arXiv:hep-ph/9607359]; E. Witten, arXiv:hep-ph/0201018; M. Dine, Y. Nir and Y. Shadmi, Phys. Rev. D **66** (2002) 115001 [arXiv:hep-ph/0206268].
12. C. Bachas, C. Fabre and T. Yanagida, Phys. Lett. B **370** (1996) 49 [arXiv:hep-th/9510094]; J. L. Chkareuli and I. G. Gogoladze, Phys. Rev. D **58** (1998) 055011 [arXiv:hep-ph/9803335].
13. V. V. Dixit and M. Sher, Phys. Rev. D **40** (1989) 3765.
14. G. R. Dvali, Phys. Lett. B **287** (1992) 101; P. Nath, Phys. Rev. Lett. **76** (1996) 2218 [arXiv:hep-ph/9512415]; P. Nath, Phys. Lett. B **381** (1996) 147 [arXiv:hep-ph/9602337]; V. Lucas and S. Raby, Phys. Rev. D **55** (1997) 6986 [arXiv:hep-ph/9610293]; Z. Berezhiani, Z. Tavartkiladze and M. Vysotsky, arXiv:hep-ph/9809301; K. Turzynski, JHEP **0210** (2002) 044 [arXiv:hep-ph/0110282]; D. Emmanuel-Costa and S. Wiesenfeldt, Nucl. Phys. B **661** (2003) 62 [arXiv:hep-ph/0302272].
15. T. Friedmann and E. Witten, arXiv:hep-th/0211269; I. R. Klebanov and E. Witten, Nucl. Phys. B **664** (2003) 3 [arXiv:hep-th/0304079].
16. R. N. Mohapatra, Phys. Rev. Lett. **43** (1979) 893.

17. S. Nandi, A. Stern and E. C. G. Sudarshan, Phys. Lett. B **113** (1982) 165; V. S. Berezinsky and A. Y. Smirnov, Phys. Lett. B **140** (1984) 49.
18. C. S. Aulakh, B. Bajc, A. Melfo, A. Rašin and G. Senjanović, Nucl. Phys. B **597** (2001) 89 [arXiv:hep-ph/0004031].
19. V. A. Rubakov, Nucl. Phys. B **203** (1982) 311; C. G. Callan, Phys. Rev. D **25** (1982) 2141 and Phys. Rev. D **26** (1982) 2058.
20. K. Freese, M. S. Turner and D. N. Schramm, Phys. Rev. Lett. **51** (1983) 1625.
21. G. Giacomelli and L. Patrizii, arXiv:hep-ex/0302011.
22. A. Melfo and G. Senjanović, Phys. Rev. D **68** (2003) 035013 [arXiv:hep-ph/0302216], and references therein.
23. R. Barbieri and L. J. Hall, Phys. Lett. B **338** (1994) 212 [arXiv:hep-ph/9408406].
24. I. Masina and C. Savoy, arXiv:hep-ph/0309067.
25. K. S. Babu and R. N. Mohapatra, Phys. Lett. B **518** (2001) 269 [arXiv:hep-ph/0108089].
26. F. Wilczek and A. Zee, Phys. Rev. D **25** (1982) 553.
27. see for example C. H. Albright and S. M. Barr, Phys. Rev. Lett. **85** (2000) 244 [arXiv:hep-ph/0002155]; K. S. Babu, J. C. Pati and F. Wilczek, Nucl. Phys. B **566** (2000) 33 [arXiv:hep-ph/9812538]; T. Blažek, S. Raby and K. Tobe, Phys. Rev. D **60** (1999) 113001 [arXiv:hep-ph/9903340].
28. M. C. Chen and K. T. Mahanthappa, arXiv:hep-ph/0305088.
29. R. N. Mohapatra, Phys. Rev. D **34** (1986) 3457.
30. T. E. Clark, T. K. Kuo and N. Nakagawa, Phys. Lett. B **115** (1982) 26; D. G. Lee, Phys. Rev. D **49** (1994) 1417.
31. C. S. Aulakh and R. N. Mohapatra, Phys. Rev. D **28** (1983) 217.
32. C. S. Aulakh, B. Bajc, A. Melfo, G. Senjanović and F. Vissani, arXiv:hep-ph/0306242.
33. M. Gell-Mann, P. Ramond and R. Slansky, proceedings of the Supergravity Stony Brook Workshop, New York, 1979, eds. P. Van Niewenhuizen and D. Freeman (North-Holland, Amsterdam); T. Yanagida, proceedings of the Workshop on Unified Theories and Baryon Number in the Universe, Tsukuba, Japan 1979 (edited by A. Sawada and A. Sugamoto, KEK Report No. 79-18, Tsukuba); R. N. Mohapatra and G. Senjanović, Phys. Rev. Lett. **44** (1980) 912.
34. M. Magg and C. Wetterich, Phys. Lett. B **94** (1980) 61; R. N. Mohapatra and G. Senjanović, Phys. Rev. D **23** (1981) 165.
35. D. G. Lee and R. N. Mohapatra, Phys. Rev. D **51** (1995) 1353 [arXiv:hep-ph/9406328].
36. K. Matsuda, Y. Koide, T. Fukuyama and H. Nishiura, Phys. Rev. D **65** (2002) 033008 [Erratum-ibid. D **65** (2002) 079904] [arXiv:hep-ph/0108202]; T. Fukuyama and N. Okada, JHEP **0211** (2002) 011 [arXiv:hep-ph/0205066].
37. B. Bajc, G. Senjanović and F. Vissani, arXiv:hep-ph/0110310.
38. B. Bajc, G. Senjanović and F. Vissani, Phys. Rev. Lett. **90** (2003) 051802 [arXiv:hep-ph/0210207].
39. H. S. Goh, R. N. Mohapatra and S. P. Ng, Phys. Lett. B **570** (2003) 215 [arXiv:hep-ph/0303055].
40. H. S. Goh, R. N. Mohapatra and S. P. Ng, arXiv:hep-ph/0308197.
41. B. Brahmachari and R. N. Mohapatra, Phys. Rev. D **58** (1998) 015001 [arXiv:hep-ph/9710371].
42. S. F. King, arXiv:hep-ph/0310204.
43. D. G. Lee and R. N. Mohapatra, Phys. Lett. B **324** (1994) 376 [arXiv:hep-ph/9310371].

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard model *(p. 150)*
Portorož, Slovenia, July 12-17, 2003

General Principles of Brane Kinematics and Dynamics

Matej Pavšič

Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
e-mail: matej.pavsic@ijs.si

Abstract. We consider branes as "points" in an infinite dimensional brane space $\mathcal{M}$ with a prescribed metric. Branes move along the geodesics of $\mathcal{M}$. For a particular choice of metric the equations of motion are equivalent to the well known equations of the Dirac-Nambu-Goto branes (including strings). Such theory describes "free fall" in $\mathcal{M}$-space. In the next step the metric of $\mathcal{M}$-space is given the dynamical role and a corresponding kinetic term is added to the action. So we obtain a background independent brane theory: a space in which branes live is $\mathcal{M}$-space and it is not given in advance, but comes out as a solution to the equations of motion. The embedding space ("target space") is not separately postulated. It is identified with the brane configuration.

1 Introduction

Theories of strings and higher dimensional extended objects, branes, are very promising in explaining the origin and interrelationship of the fundamental interactions, including gravity. But there is a cloud. It is not clear what is a geometric principle behind string and brane theories and how to formulate them in a background independent way. An example of a background independent theory is general relativity where there is no preexisting space in which the theory is formulated. The dynamics of the 4-dimensional space (spacetime) itself results as a solution to the equations of motion. The situation is sketched in Fig.1. A point particle traces a world line in spacetime whose dynamics is governed by the Einstein-Hilbert action. A closed string traces a world tube, but so far its has not been clear what is the appropriate space and action for a background independent formulation of string theory.

Here I will report about a formulation of string and brane theory (see also ref. [1]) which is based on the infinite dimensional brane space $\mathcal{M}$. The "points" of this space are branes and their coordinates are the brane (embedding) functions. In $\mathcal{M}$-space we can define the distance, metric, connection, covariant derivative, curvature, etc. We show that the brane dynamics can be derived from the principle of minimal length in $\mathcal{M}$-space; a brane follows a geodetic path in $\mathcal{M}$. The situation is analogous to the free fall of an ordinary point particle as described by general relativity. Instead of keeping the metric fixed, we then add to the action a kinetic term for the metric of $\mathcal{M}$-space and so we obtain a background independent brane theory in which there is no preexisting space.

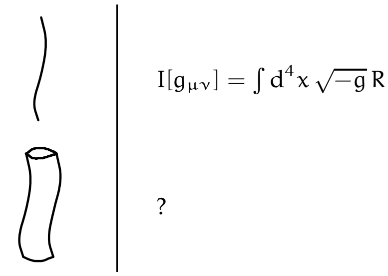

Fig. 1. To point particle there corresponds the Einstein-Hilbert action in spacetime. What is a corresponding space and action for a closed string?

2 Brane space $\mathcal{M}$ (brane kinematics)

We will first treat the brane kinematics, and only later we will introduce a brane dynamics. We assume that the basic kinematically possible objects are n-dimensional, arbitrarily deformable branes $\mathcal{V}_n$ living in an N-dimensional embedding (target) space V_N. Tangential deformations are also allowed. This is illustrated in Fig. 2. Imagine a rubber sheet on which we paint a grid of lines. Then we deform the sheet in such a way that mathematically the surface remains the same, nevertheless the deformed object is physically different from the original object.

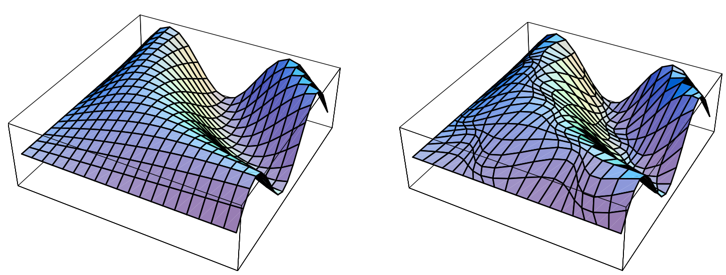

Fig. 2. Examples of tangentially deformed membranes. Mathematically the surface on the left and is the same as the surface on the right. Physically the two surfaces are different.

We represent $\mathcal{V}_n$ by functions $X^\mu(\xi^a)$, $\mu = 0, 1, ..., N-1$, where ξ^a, $a = 0, 1, 2, ..., n-1$ are parameters on $\mathcal{V}_n$. According the assumed interpretation, different functions $X^\mu(\xi^a)$ can represent physically different branes. That is, if we perform an *active diffeomorphism* $\xi^a \rightarrow \xi'^a = f^a(\xi)$, then the new functions $X^\mu(f^a(\xi)) = X'^\mu(\xi)$ represent a physically different brane $\mathcal{V}'_n$. For a more complete and detailed discussion see ref. [1].

The set of all possible $\mathcal{V}_n$ forms *the brane space* $\mathcal{M}$. A brane $\mathcal{V}_n$ can be considered as a point in $\mathcal{M}$ parametrized by coordinates $X^\mu(\xi^a) \equiv X^{\mu(\xi)}$ which bear

a discrete index μ and n continuous indices ξ^a. That is, $\mu(\xi)$ as superscript or subscript denotes a single index which consists of the discrete part μ and the continuous part (ξ).

In analogy with the finite-dimensional case we can introduce the *distance* $\mathrm{d}\ell$ in the infinite-dimensional space $\mathcal{M}$:

$$\mathrm{d}\ell^2 = \int \mathrm{d}\xi\, \mathrm{d}\zeta\, \rho_{\mu\nu}(\xi,\zeta)\, \mathrm{d}X^\mu(\xi)\, \mathrm{d}X^\nu(\zeta) = \rho_{\mu(\xi)\nu(\zeta)}\, \mathrm{d}X^{\mu(\xi)}\, \mathrm{d}X^{\nu(\zeta)},$$

where $\rho_{\mu\nu}(\xi,\zeta) \equiv \rho_{\mu(\xi)\nu(\zeta)}$ is the metric in $\mathcal{M}$. Let us consider a particular choice of metric

$$\rho_{\mu(\xi)\nu(\zeta)} = \sqrt{|f|}\, \alpha\, g_{\mu\nu} \delta(\xi - \zeta), \tag{1}$$

where $f \equiv \det f_{ab}$ is the determinant of the induced metric $f_{ab} \equiv \partial_a X^\alpha \partial_b X^\beta\, g_{\alpha\beta}$ on the sheet V_n, whilst $g_{\mu\nu}$ is the metric tensor of the embedding space V_N, and α an arbitrary function of ξ^a or, in particular, a constant. Then the line element (1) becomes

$$\mathrm{d}\ell^2 = \int \mathrm{d}\xi \sqrt{|f|}\, \alpha\, g_{\mu\nu}\, \mathrm{d}X^\mu(\xi) \mathrm{d}X^\nu(\xi). \tag{2}$$

The invariant volume (measure) in $\mathcal{M}$ is

$$\sqrt{|\rho|}\mathcal{D}X = (\mathrm{Det}\, \rho_{\mu\nu}(\xi,\zeta))^{1/2} \prod_{\xi,\mu} \mathrm{d}X^\mu(\xi). \tag{3}$$

Here Det denotes a continuum determinant taken over ξ, ζ as well as over μ, ν. In the case of the diagonal metric (1) we have

$$\sqrt{|\rho|}\mathcal{D}X = \prod_{\xi,\mu} \left(\sqrt{|f|}\, \alpha\, |g|\right)^{1/2} \mathrm{d}X^\mu(\xi) \tag{4}$$

Tensor calculus in $\mathcal{M}$-space is analogous to that in a finite dimensional space. The differential of coordinates $\mathrm{d}X^\mu(\xi) \equiv \mathrm{d}X^{\mu(\xi)}$ is a vector in $\mathcal{M}$. The coordinates $X^{\mu(\xi)}$ can be transformed into new coordinates $X'^{\mu(\xi)}$ which are functionals of $X^{\mu(\xi)}$:

$$X'^{\mu(\xi)} = F^{\mu(\xi)}[X]. \tag{5}$$

If functions $X^\mu(\xi)$ represent a brane $\mathcal{V}_n$, then functions $X'^\mu(\xi)$ obtained from $X^\mu(\xi)$ by a functional transformation represent the same (kinematically possible) brane.

Under a general coordinate transformation (5) a generic vector $A^{\mu(\xi)} \equiv A^\mu(\xi)$ transforms as[1]

$$A^{\mu(\xi)} = \frac{\partial X'^{\mu(\xi)}}{\partial X^{\nu(\zeta)}} A^{\nu(\zeta)} \equiv \int \mathrm{d}\zeta \frac{\delta X'^\mu(\xi)}{\delta X^\nu(\zeta)} A^\nu(\zeta)\,, \tag{6}$$

where $\delta/\delta X^\mu(\xi)$ denotes the functional derivative.

[1] A similar formalism, but for a specific type of the functional transformations, namely the reparametrizations which functionally depend on string coordinates, was developed by Bardakci [2]

Similar transformations hold for a covariant vector $A_{\mu(\xi)}$, a tensor $B_{\mu(\xi)\nu(\zeta)}$, etc.. Indices are lowered and raised, respectively, by $\rho_{\mu(\xi)\nu(\zeta)}$ and $\rho^{\mu(\xi)\nu(\zeta)}$, the latter being the inverse metric tensor satisfying

$$\rho^{\mu(\xi)\alpha(\eta)}\rho_{\alpha(\eta)\nu(\zeta)} = \delta^{\mu(\xi)}{}_{\nu(\zeta)}. \tag{7}$$

As can be done in a finite-dimensional space, we can also define the *covariant derivative* in $\mathcal{M}$. When acting on a *scalar* $A[X(\xi)]$ the covariant derivative coincides with the ordinary functional derivative:

$$A_{;\mu(\xi)} = \frac{\delta A}{\delta X^{\mu}(\xi)} \equiv A_{,\mu(\xi)}. \tag{8}$$

But in general a geometric object in $\mathcal{M}$ is a tensor of arbitrary rank,

$$A^{\mu_1(\xi_1)\mu_2(\xi_2)\dots}{}_{\nu_1(\zeta_1)\nu_2(\zeta_2)\dots},$$

which is a functional of $X^{\mu}(\xi)$, and its covariant derivative contains the affinity $\Gamma^{\mu(\xi)}_{\nu(\zeta)\sigma(\eta)}$ composed of the metric $\rho_{\mu(\xi)\nu(\xi')}$ [3]. For instance, when acting on a vector the covariant derivative gives

$$A^{\mu(\xi)}{}_{;\nu(\zeta)} = A^{\mu(\xi)}{}_{,\nu(\zeta)} + \Gamma^{\mu(\xi)}_{\nu(\zeta)\sigma(\eta)}A^{\sigma(\eta)} \tag{9}$$

In a similar way we can write the covariant derivative acting on a tensor of arbitrary rank.

In analogy to the notation as employed in the finite dimensional tensor calculus we can use the following variants of notation for the ordinary and covariant derivative:

$$\begin{aligned} &\frac{\delta}{\delta X^{\mu}(\xi)} \equiv \frac{\partial}{\partial X^{\mu(\xi)}} \equiv \partial_{\mu(\xi)} \quad \text{for functional derivative} \\ &\frac{D}{DX^{\mu}(\xi)} \equiv \frac{D}{DX^{\mu(\xi)}} \equiv D_{\mu(\xi)} \quad \text{for covariant derivative in } \mathcal{M} \end{aligned} \tag{10}$$

Such shorthand notations for functional derivative is very effective.

3 Brane dynamics: brane theory as free fall in $\mathcal{M}$-space

So far we have considered kinematically possible branes as the points in the brane space $\mathcal{M}$. Instead of one brane we can consider a one parameter family of branes $X^{\mu}(\tau, \xi^{a}) \equiv X^{\mu(\xi)}(\tau)$, i.e., a curve (or trajectory) in $\mathcal{M}$. Every trajectory is kinematically possible in principle. A particular dynamical theory then selects which amongst those kinematically possible branes and trajectories are also dynamically possible. We will assume that dynamically possible trajectories are *geodesics* in $\mathcal{M}$ described by the minimal length action [1]:

$$I[X^{\alpha(\xi)}] = \int d\tau' \left(\rho_{\alpha(\xi')\beta(\xi'')}\dot{X}^{\alpha(\xi')}\dot{X}^{\beta(\xi'')}\right)^{1/2}. \tag{11}$$

Let us introduce the shorthand notation

$$\mu \equiv \rho_{\alpha(\xi')\beta(\xi'')} \dot{X}^{\alpha(\xi')} \dot{X}^{\beta(\xi'')} \tag{12}$$

and vary the action (11) with respect to $X^{\alpha(\xi)}(\tau)$. If the expression for the metric $\rho_{\alpha(\xi')\beta(\xi'')}$ does not contain the velocity $\dot{X}^\mu$ we obtain

$$\frac{1}{\mu^{1/2}} \frac{d}{d\tau} \left(\frac{\dot{X}^{\mu(\xi)}}{\mu^{1/2}} \right) + \Gamma^{\mu(\xi)}{}_{\alpha(\xi')\beta(\xi'')} \frac{\dot{X}^{\alpha(\xi')} \dot{X}^{\beta(\xi'')}}{\mu} = 0 \tag{13}$$

which is a straightforward generalization of the usual geodesic equation from a finite-dimensional space to an infinite-dimensional $\mathcal{M}$-space.

Let us now consider a particular choice of the $\mathcal{M}$-space metric:

$$\rho_{\alpha(\xi')\beta(\xi'')} = \kappa \frac{\sqrt{|f(\xi')|}}{\sqrt{\dot{X}^2(\xi')}} \delta(\xi' - \xi'') \eta_{\alpha\beta} \tag{14}$$

where $\dot{X}^2 \equiv g_{\mu\nu} \dot{X}^\mu \dot{X}^\nu$ is the square of velocity $\dot{X}^\mu$. Therefore, the metric (14) depends on velocity. If we insert it into the action (11), then after performing the functional derivatives and the integrations over τ and ξ^a (implied in the repeated indexes $\alpha(\xi')$, $\beta(\xi'')$) we obtain the following equations of motion:

$$\frac{d}{d\tau} \left(\frac{1}{\mu^{1/2}} \frac{\sqrt{|f|}}{\sqrt{\dot{X}^2}} \dot{X}_\mu \right) + \frac{1}{\mu^{1/2}} \partial_a \left(\sqrt{|f|} \sqrt{\dot{X}^2} \partial^a X_\mu \right) = 0 \tag{15}$$

If we take into account the relations

$$\frac{d\sqrt{|f|}}{d\tau} = \frac{\partial \sqrt{|f|}}{\partial f_{ab}} \dot{f}_{ab} = \sqrt{|f|}\, f^{ab} \partial_a \dot{X}^\mu \partial_b X_\mu = \sqrt{|f|}\, \partial^a X_\mu \partial_a \dot{X}^\mu \tag{16}$$

and

$$\frac{\dot{X}_\mu}{\sqrt{\dot{X}^2}} \frac{\dot{X}^\mu}{\sqrt{\dot{X}^2}} = 1 \quad \Rightarrow \quad \frac{d}{d\tau} \left(\frac{\dot{X}_\mu}{\sqrt{\dot{X}^2}} \right) \dot{X}^\mu = 0 \tag{17}$$

it is not difficult to find that

$$\frac{d\mu}{d\tau} = 0 \tag{18}$$

Therefore, instead of (15) we can write

$$\frac{d}{d\tau} \left(\frac{\sqrt{|f|}}{\sqrt{\dot{X}^2}} \dot{X}_\mu \right) + \partial_a \left(\sqrt{|f|} \sqrt{\dot{X}^2} \partial^a X_\mu \right) = 0. \tag{19}$$

This are precisely the equation of motion for the Dirac-Nambu-Goto brane, written in a particular gauge.

The action (11) is by definition invariant under reparametrizations of ξ^a. In general, it is not invariant under reparametrization of the parameter τ. If the expression for the metric $\rho_{\alpha(\xi')\beta(\xi'')}$ does not contain the velocity $\dot{X}^\mu$, then the action (11) is invariant under reparametrizations of τ. This is no longer true if $\rho_{\alpha(\xi')\beta(\xi'')}$ contains $\dot{X}^\mu$. Then the action (11) is not invariant under reparametrizations of τ.

In particular, if metric is given by eq. (14), then the action becomes explicitly

$$I[X^{\mu}(\xi)] = \int d\tau \left(d\xi \, \kappa \sqrt{|f|} \sqrt{\dot{X}^2} \right)^{1/2} \tag{20}$$

and the equations of motion (15), as we have seen, automatically contain the relation

$$\frac{d}{d\tau} \left(\dot{X}^{\mu(\xi)} \dot{X}_{\mu(\xi)} \right) \equiv \frac{d}{d\tau} \int d\xi \, \kappa \sqrt{|f|} \sqrt{\dot{X}^2} = 0. \tag{21}$$

The latter relation is nothing but *a gauge fixing relation*, where by "gauge" we mean here a choice of parameter τ. The action (11), which in the case of the metric (14) is not reparametrization invariant, contains the gauge fixing term.

In general the exponent in the Lagrangian is not necessarily $\frac{1}{2}$, but can be arbitrary:

$$I[X^{\alpha(\xi)}] = \int d\tau \left(\rho_{\alpha(\xi')\beta(\xi'')} \dot{X}^{\alpha(\xi')} \dot{X}^{\beta(\xi'')} \right)^{a}. \tag{22}$$

For the metric (14) we have explicitly

$$I[X^{\mu}(\xi)] = \int d\tau \left(d\xi \, \kappa \sqrt{|f|} \sqrt{\dot{X}^2} \right)^{a} \tag{23}$$

The corresponding equations of motion are

$$\frac{d}{d\tau} \left(a \mu^{a-1} \frac{\kappa \sqrt{|f|}}{\sqrt{\dot{X}^2}} \dot{X}_{\mu} \right) + a \mu^{a-1} \partial_a \left(\kappa \sqrt{|f|} \sqrt{\dot{X}^2} \partial^a X_{\mu} \right) = 0. \tag{24}$$

We distinguish two cases:

(i) $a \neq 1$. Then the action is *not* invariant under reparametrizations of τ. The equations of motion (24) for $a \neq 1$ imply the gauge fixing relation $d\mu/d\tau = 0$, that is, the relation (21).

(ii) $a = 1$. Then the action (23) is invariant under reparametrizations of τ. The equations of motion for $a = 1$ contain no gauge fixing term. In both cases, (i) and (ii), we obtain the same equations of motion (19).

Let us focus our attention to the action with $a = 1$:

$$I[X^{\alpha(\xi)}] = \int d\tau \left(\rho_{\alpha(\xi')\beta(\xi'')} \dot{X}^{\alpha(\xi')} \dot{X}^{\beta(\xi'')} \right) = \int d\tau \, d\xi \, \kappa \sqrt{|f|} \sqrt{\dot{X}^2} \tag{25}$$

It is invariant under the transformations

$$\tau \to \tau' = \tau'(\tau) \tag{26}$$

$$\xi^a \to \xi'^a = \xi'^a(\xi^a) \tag{27}$$

in which τ and ξ^a do not mix.

Invariance of the action (25) under reparametrizations (26) of the evolution parameter τ implies the existence of a constraint among the canonical momenta $p_{\mu(\xi)}$ and coordinates $X^{\mu(\xi)}$. Momenta are given by

$$\begin{aligned} p_{\mu(\xi)} &= \frac{\partial L}{\partial \dot{X}^{\mu(\xi)}} = 2\rho_{\mu(\xi)\nu(\xi')} \dot{X}^{\nu(\xi')} + \frac{\partial \rho_{\alpha(\xi')\beta(\xi'')}}{\partial \dot{X}^{\mu(\xi)}} \dot{X}^{\alpha(\xi')} \dot{X}^{\beta(\xi'')} \\ &= \frac{\kappa \sqrt{|f|}}{\sqrt{\dot{X}^2}} \dot{X}_{\mu}. \end{aligned} \tag{28}$$

By distinguishing covariant and contravariant components one finds

$$p_{\mu(\xi)} = \dot{X}_{\mu(\xi)} = \rho_{\mu(\xi)\nu(\xi')}\dot{X}^{\nu(\xi')} \ , \quad p^{\mu(\xi)} = \dot{X}^{\mu(\xi)}. \tag{29}$$

We define $p_{\mu(\xi)} \equiv p_\mu(\xi) \equiv p_\mu$, $\dot{X}^{\mu(\xi)} \equiv \dot{X}^\mu(\xi) \equiv \dot{X}^\mu$. Here p_μ and $\dot{X}^\mu$ have the meaning of the usual finite dimensional vectors whose components are lowered and raised by the finite-dimensional metric tensor $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$: $p^\mu = g^{\mu\nu}p_\nu$, $\dot{X}_\mu = g_{\mu\nu}\dot{X}^\nu$.

The *Hamiltonian* belonging to the action (25) is

$$H = p_{\mu(\xi)}\dot{X}^{\mu(\xi)} - L = \int d\xi \, \frac{\sqrt{\dot{X}^2}}{\kappa\sqrt{|f|}} \left(p^\mu p_\mu - \kappa^2|f|\right) = p_{\mu(\xi)}p^{\mu(\xi)} - K = 0 \tag{30}$$

where $K = K[X^{\mu(\xi)}] = \int d\xi \, \kappa \sqrt{|f|}\sqrt{\dot{X}^2} = L$. It is identically zero. The $\dot{X}^2$ entering the integral for H is arbitrary due to arbitrary reparametrizations of τ (which change $\dot{X}^2$). Consequently, H vanishes when the following expression under the integral vanishes:

$$p^\mu p_\mu - \kappa^2|f| = 0 \tag{31}$$

Expression (31) is the usual constraint for the Dirac-Nambu-Goto brane (p-brane). It is satisfied at every ξ^a.

In ref. [1] it is shown that the constraint is conserved in τ and that as a consequence we have

$$p_\mu \partial_a X^\mu = 0. \tag{32}$$

The latter equation is yet another set of constraints[2] which are satisfied at any point ξ^a of the brane world manifold V_{n+1}.

Both kinds of constraints are thus automatically implied by the action (25) in which the choice (14) of $\mathcal{M}$-space metric tensor has been taken.

Introducing a more compact notation $\phi^A = (\tau, \xi^a)$ and $X^{\mu(\xi)}(\tau) \equiv X^\mu(\phi^A) \equiv X^{\mu(\phi)}$ we can write

$$I[X^{\mu(\phi)}] = \rho_{\mu(\phi)\nu(\phi')}\dot{X}^{\mu(\phi)}\dot{X}^{\nu(\phi')} = \int d^{n+1}\phi \, \sqrt{|f|}\,\sqrt{\dot{X}^2} \tag{33}$$

where

$$\rho_{\mu(\phi')\nu(\phi'')} = \kappa \frac{\sqrt{|f(\xi')|}}{\sqrt{\dot{X}^2(\xi')}} \, \delta(\xi' - \xi'')\delta(\tau' - \tau'')\eta_{\mu\nu} \tag{34}$$

Variation of the action (33) with respect to $X^{\mu(\phi)}$ gives

$$\frac{d\dot{X}^{\mu(\phi)}}{d\tau} + \Gamma^{\mu(\phi)}_{\alpha(\phi')\beta(\phi'')}\dot{X}^{\alpha(\phi')}\dot{X}^{\beta(\phi'')} = 0 \tag{35}$$

which is the geodesic equation in the space $\mathcal{M}_{V_{n+1}}$ of brane world manifolds V_{n+1} described by $X^{\mu(\phi)}$. For simplicity we will omit the subscript and call the latter space $\mathcal{M}$-space as well.

[2] Something similar happens in canonical gravity. Moncrief and Teitelboim [4] have shown that if one imposes the Hamiltonian constraint on the Hamilton functional then the momentum constraints are automatically satisfied.

Once we have the constraints we can write the first order or phace space action

$$I[X^\mu, p_\mu, \lambda, \lambda^a] = \int d\tau\, d\xi \left(p_\mu \dot{X}^\mu - \frac{\lambda}{2\kappa\sqrt{|f|}}(p^\mu p_\mu - \kappa^2|f|) - \lambda^a p_\mu \partial_a X^\mu \right), \quad (36)$$

where λ and λ^a are Lagrange multipliers. It is classically equivalent to the *minimal surface action* for the $(n+1)$-dimensional world manifold V_{n+1}

$$I[X^\mu] = \kappa \int d^{n+1}\phi\, (\det \partial_A X^\mu \partial_B X_\mu)^{1/2}. \quad (37)$$

This is the conventional Dirac–Nambu–Goto action, invariant under reparametrizations of ϕ^A.

4 Dynamical metric field in $\mathcal{M}$-space

Let us now ascribe the dynamical role to the $\mathcal{M}$-space metric. ¿From $\mathcal{M}$-space perspective we have motion of a point "particle" in the presence of a metric field $\rho_{\mu(\phi)\nu(\phi')}$ which is itself dynamical.

As a model let us consider the action

$$I[\rho] = \int \mathcal{D}X\sqrt{|\rho|} \left(\rho_{\mu(\phi)\nu(\phi')}\dot{X}^{\mu(\phi)}\dot{X}^{\nu(\phi')} + \frac{\epsilon}{16\pi}\mathcal{R} \right). \quad (38)$$

where ρ is the determinant of the metric $\rho_{\mu(\phi)\nu(\phi')}$ and ϵ a constant. Here $\mathcal{R}$ is the Ricci scalar in $\mathcal{M}$-space, defined according to $\mathcal{R} = \rho^{\mu(\phi)\nu(\phi')}\mathcal{R}_{\mu(\phi)\nu(\phi')}$, where $\mathcal{R}_{\mu(\phi)\nu(\phi')}$ is the Ricci tensor in $\mathcal{M}$-space [1].

Variation of the action (38) with respect to $X^{\mu(\phi)}$ and $\rho_{\mu(\phi)\nu(\phi')}$ leads to (see ref.[1]) the *geodesic equation* (35) and to the *Einstein equations* in $\mathcal{M}$-space

$$\dot{X}^{\mu(\phi)}\dot{X}^{\nu(\phi)} + \frac{\epsilon}{16\pi}\mathcal{R}^{\mu(\phi)\nu(\phi')} = 0 \quad (39)$$

In fact, after performing the variation we had a term with $\mathcal{R}$ and a term with $\dot{X}^{\mu(\phi)}\dot{X}_{\mu(\phi)}$ in the Einstein equations. But, after performing the contraction with the metric, we find that the two terms cancel each other resulting in the simplified equations (39) (see ref.[1]).

The metric $\rho_{\mu(\phi)\nu(\phi')}$ is a functional of the variables $X^{\mu(\phi)}$ and in eqs. (35),(39) we have a system of functional differential equations which determine the set of possible solutions for $X^{\mu(\phi)}$ and $\rho_{\mu(\phi)\nu(\phi')}$. Our brane model (including strings) is background independent: there is no preexisting space with a preexisting metric, neither curved nor flat.

We can imagine a model universe consisting of a single brane. Although we started from a brane embedded in a higher dimensional finite space, we have subsequently arrived at the action(38) in which the dynamical variables $X^{\mu(\phi)}$ and $\rho_{\mu(\phi)\nu(\phi')}$ are defined in $\mathcal{M}$-space. In the latter model the concept of an underlying finite dimensional space, into which the brane is embedded, is in fact

abolished. We keep on talking about "branes" for convenience reasons, but actually there is no embedding space in this model. The metric $\rho_{\mu(\phi)\nu(\phi')}[X]$ is defined only on the brane. There is no metric of a space into which the brane is embedded. Actually, there is no embedding space. All what exists is a brane configuration $X^{\mu(\phi)}$ and the corresponding metric $\rho_{\mu(\phi)\nu(\phi')}$ in $\mathcal{M}$-space.

A system of branes (a brane configuration) Instead of a single brane we can consider a system of branes described by coordinates $X^{\mu(\phi,k)}$, where k is a discrete index that labels the branes (Fig. 3). If we replace (ϕ) with (ϕ, k), or, alternatively, if we interpret (ϕ) to include the index k, then the previous action (38) and equations of motion (35),(39) are also valid for a system of branes.

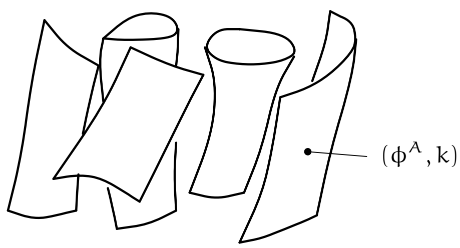

Fig. 3. The system of branes is represented as being embedded in a finite-dimensional space V_N. The concept of a continuous embedding space is only an approximation which, when there are many branes, becomes good at large scales (i.e., at the "macroscopic" level). The metric is defined only at the points (ϕ, k) situated on the branes. At large scales (or equivalently, when the branes are "small" and densely packed together) the set of all the points (ϕ, k) is a good approximation to a continuous metric space V_N.

A brane configuration is all what exists in such a model. It is identified with the embedding space[3].

From $\mathcal{M}$-space to spacetime We now define $\mathcal{M}$-space as the space of all possible brane configurations. Each brane configuration is considered as a point in $\mathcal{M}$-space described by coordinates $X^{\mu(\phi,k)}$. The metric $\rho_{\mu(\phi,k)\nu(\phi',k')}$ determines the *distance* between two points *belonging to two different brane configurations*:

$$d\ell^2 = \rho_{\mu(\phi,k)\nu(\phi',k')} dX^{\mu(\phi,k)} dX^{\nu(\phi',k')} \tag{40}$$

[3] Other authors also considered a class of brane theories in which the embedding space has no prior existence, but is instead coded completely in the degrees of freedom that reside on the branes. They take for granted that, as the background is not assumed to exist, there are no embedding coordinates (see e.g., [5]). This seems to be consistent with our usage of $X^{\mu(\phi)}$ which, at the fundamental level, are not considered as the embedding coordinates, but as the $\mathcal{M}$-space coordinates. Points of $\mathcal{M}$-space are described by coordinates $X^{\mu(\phi)}$, and the distance between the points is determined by the metric $\rho_{\mu(\phi)\nu(\phi')}$, which is dynamical.. In the limit of infinitely many densely packed branes, the set of points (ϕ^A, k) is supposed to become a continuous, finite dimensional metric space V_N.

where

$$dX^{\mu(\phi,k)} = X'^{\mu(\phi,k)} - X^{\mu(\phi,k)}. \tag{41}$$

Let us now introduce another quantity which connects two different points, in the usual sense of the word, *within the same brane configuration*:

$$\widetilde{\Delta} X^{\mu}(\phi, k) \equiv X^{\mu(\phi',k')} - X^{\mu(\phi,k)}. \tag{42}$$

and define

$$\Delta s^2 = \rho_{\mu(\phi,k)\nu(\phi',k')} \widetilde{\Delta} X^{\mu}(\phi, k) \widetilde{\Delta} X^{\nu}(\phi', k'). \tag{43}$$

In the above formula summation over the repeated indices μ and ν is assumed, but no integration over ϕ, ϕ' and no summation over k, k'.

Eq.(43) denotes the distance between the points within a given brane configuration. This is the quadratic form in the skeleton space S. The metric ρ in the skeleton space S is the prototype of the metric in target space V_N (the embedding space). A brane configuration is a skeleton S of a target space V_N.

5 Conclusion

We have taken the brane space $\mathcal{M}$ seriously as an arena for physics. The arena itself is also a part of the dynamical system, it is not prescribed in advance. The theory is thus background independent. It is based on a geometric principle which has its roots in the brane space $\mathcal{M}$. We can thus complete the picture that occurred in the introduction:

$$I[g_{\mu\nu}] = \int d^4x \sqrt{-g}\, R$$

$$I[\rho_{\mu(\phi)\nu(\phi')}] = \int \mathcal{D}X \sqrt{|\rho|}\, \mathcal{R}$$

Fig. 4. Brane theory is formulated in $\mathcal{M}$-space. The action is given in terms of the $\mathcal{M}$-space curvature scalar $\mathcal{R}$.

We have formulated a theory in which an embedding space *per se* does not exist, but is intimately connected to the existence of branes (including strings). Without branes there is no embedding space. There is no preexisting space and metric: they appear dynamically as solutions to the equations of motion. Therefore the model is background independent.

All this was just an introduction into a generalized theory of branes. Much more can be found in a book [1] where the description with a metric tensor has been surpassed. Very promising is the description in terms of the Clifford algebra

equivalent of the tetrad which simplifies calculations significantly. The relevance of the concept of Clifford space for physics is discussed in refs. [1], [6]–[10]).

There are possible connections to other topics. The system, or condensate of branes (which, in particular, may be so dense that the corresponding points form a continuum), represents a *reference system* or *reference fluid* with respect to which the points of the target space are defined. Such a system was postulated by DeWitt [11], and recently reconsidered by Rovelli [12] in relation to the famous Einstein's 'hole argument' according to which the points of spacetime cannot be identified. The brane model presented here can also be related to the *Mach principle* according to which the motion of matter at a given location is determined by the contribution of all the matter in the universe and this provides an explanation for inertia (and inertial mass). Such a situation is implemented in the model of a universe consisting of a system of branes described by eqs. (35),(39): the motion of a k-th brane, including its inertia (metric), is determined by the presence of all the other branes.

Acknowledgement

This work has been supported by the Ministry of Education, Science and Sport of Slovenia under the contract PO-0517.

References

1. M. Pavšič, *The Landscape of Theoretical Physics: A Global View. From Point Particles to the Brane World and Beyond, in Search of a Unifying Principle*, Kluwer, Dordrecht, 2001.
2. K. Bardakci, Nucl. Phys. **B271** (1986) 561.
3. M. Pavšič, Found. Phys. **26** (1996) 159; Nuov. Cim. A **110** (1997) 369.
4. V. Moncrief and C. Teitelboim, Phys. Rev. D **6** (1972) 966.
5. L. Smolin, Phys. Rev. D **62** (2000) 086001.
6. C. Castro, J. Chaos Solitons Fractals **11** (2000) 1721, hep-th/9912113; J. Chaos, Solitons and Fractals **12** (2001) 1585, physics/0011040.
7. W.M. Pezzaglia Jr, *A Clifford Algebra Multivector Reformulation of Field Theory*, Dissertation, University of California, Davis, 1983; "Classification of Multivector Theories and Modification o f the Postulates of Physics", gr-qc/9306006; "Polydimensional Relativity, a Classical Generalization of the Automorphism Invariance Principle", gr-qc/9608052; "Physical Applications of a Generalized Clifford Calculus: Papapetrou Equations and Metamorphic Curvature", gr-qc/9710027.
8. M. Pavšič, *Found. Phys.* **31**, 1185 (2001) 1185, hep-th/0011216; Found. Phys. **33** (2003) 1277, gr-qc/0211085; Class. Quant. Grav. **20** (2003) 2697, gr-qc/0111092; C. Castro and M. Pavšič, Phys. Lett. B **539** (2002) 133, hep-th/0110079.
9. D. Hestenes, *Space-Time Algebra*, Gordon and Breach, New York, 1966; D. Hestenes *Clifford Algebra to Geometric Calculus* D. Reidel, Dordrecht, 1984.
10. N. S. Mankoč Borštnik and H. B. Nielsen, J. Math. Phys. **43** (2002) 5782, hep-th/0111257; J. Math. Phys. **44** (2003) 4817, hep-th/0303224.
11. B.S. DeWitt, The Quantization of Geometry, in: L. Witten, *Gravitation: An Introduction to Current Research*, Wiley, New York, 1962.
12. C. Rovelli, Classical and Quantum Gravity **8** (1991) 297; **8** (1991) 317.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 161)*
Portorož, Slovenia, July 12-17, 2003

Cosmological Neutrinos

Gianpiero Mangano

INFN, Sezione di Napoli,
Dipartimento di Scienze Fisiche, Universita' di Napoli Federico II, Italy

Abstract. We review present information on cosmological neutrinos, and more generally on relativistic degrees of freedom at the Cosmic Microwave Background formation epoch, in view of the recent results of WMAP collaborations on temperature anisotropies of the CMB, as well as of recent detailed analysis of Primordial Nucleosynthesis.

1 Introduction

We are pretty confident that our Universe is presently filled with quite a large number of neutrinos, of the order of $100\ \mathrm{cm}^{-3}$ for each flavor, despite of the fact that there are no direct evidences for this claim and, more sadly, it will be also very hard to achieve this goal in the future. However several stages of the evolution of the Universe have been influenced by neutrinos, and their $silent$ contribution has been first communicated to other species via weak interactions, and eventually through their coupling with gravity. In fact, Big Bang Nucleosynthesis (BBN), the Cosmic Microwave Background (CMB) and the spectrum of Large Scale Structure (LSS) keep traces of their presence, so that by observing the power spectrum $P(k)$, the photon temperature-temperature angular correlation, and primordial abundances of light nuclei, we can obtain important pieces of information on several features of the neutrino background, as well as on some fundamental parameters, such as their mass scale. It is astonishing, at least for all those of the elementary particle community who moved to "astroparticle" physics, to see that in fact the present bound on the neutrino mass , order 1 eV, obtained by studying their effect on suppressing structure formation at small scales, is already stronger than the limit obtained in terrestrial measurement from ^{3}H decay.

In this short lecture I briefly review some of the cosmological observables which indeed lead to relevant information on both dynamical (number density, chemical potential) and kinematical (masses) neutrino properties, as well as on extra weakly coupled light species.

2 Cosmological neutrinos: standard features

For large temperatures neutrinos are kept in thermodynamical equilibrium with other species, namely $e^- - e^+$ and nucleons, which in turn share the very same temperature of photons because of electromagnetic interactions. The key

phenomenon for cosmological neutrinos is that for temperatures of the order of $T_d \sim 1MeV$ weak interactions become unable to sustain equilibrium, since the corresponding effective rate Γ_w (cross-section σ_w times electron number density n_e) falls below the expansion rate H, the Hubble parameter. We can in fact estimate $\sigma_w \sim G_F^2 T^2$ and $n_e \sim T^3$, so that $\Gamma_w = G_F^2 T^5$, and since for a radiation dominated universe $H \sim \sqrt{G_N} T^2$, with G_N the Newton constant[1] it is straightforward to get

$$T_d \sim \frac{G_N^{1/6}}{G_F^{2/3}} \sim 1MeV \tag{1}$$

From this epoch on, neutrinos freely stream with an (almost) perfect Fermi-Dirac distribution, the one they had at decoupling, while momentum red-shifts as expansion goes on. In terms of the comoving momentum $y = ka$, with a the scale factor,

$$dn_\nu = a^{-3} \frac{1}{e^y + 1} \frac{d^3 y}{(2\pi)^3} \tag{2}$$

Actually neutrinos are slightly heated up during the $e^- - e^+$ annihilation phase, which takes place at temperatures of the order of the electron mass and release entropy mainly to photons, but also to neutrinos. This is because the neutrino decoupling is not an instantaneous phenomenon, but it partially overlaps the $e^- - e^+$ annihilation phase. A detailed analysis, which also takes into account QED plasma effects on the $e^- - e^+$ pairs [1] shows that neutrino distribution is slightly different than a pure black body distribution, and the corresponding energy density differ from the instantaneous decoupling result at the level of percent.

It is customary to parameterize the contribution ρ_R of relativistic species to the expansion rate of the universe in terms of the $effective$ neutrino number N_{eff} defined as follows

$$\rho_R = \rho_\gamma + \rho_\nu + \rho_X = \left[1 + \frac{7}{8} \left(\frac{4}{11} \right)^{4/3} N_{eff} \right] \rho_\gamma \tag{3}$$

with ρ_γ, ρ_ν and ρ_X the energy density of photons, neutrinos and of extra (unknown) light species, respectively. The two factors $7/8$ and $(4/11)^{4/3}$ are due to the difference in the form of equilibrium distribution (Bose-Einstein for photons, Fermi-Dirac for neutrinos) and to the different temperature of photons and neutrinos after $e^+ - e^-$ pair annihilation. With this definition, three massless neutrinos with a pure equilibrium distribution and zero chemical potential give $N_{eff} = 3$. In view of the partial neutrino reheating from $e^+ - e^-$ the actual value is slightly larger $N_{eff} = 3.04$. The role of this parameter is crucial in our understanding of fundamental physics. Any result in favor of a larger (or a smaller) value for N_{eff} would imply some exotic non standard physics at work in the early universe. In the following Sections we will see how this parameter is in fact constrained by some crucial cosmological observables, such as CMB or BBN.

[1] I adopt the standard unit system $\hbar = c = k = 1$

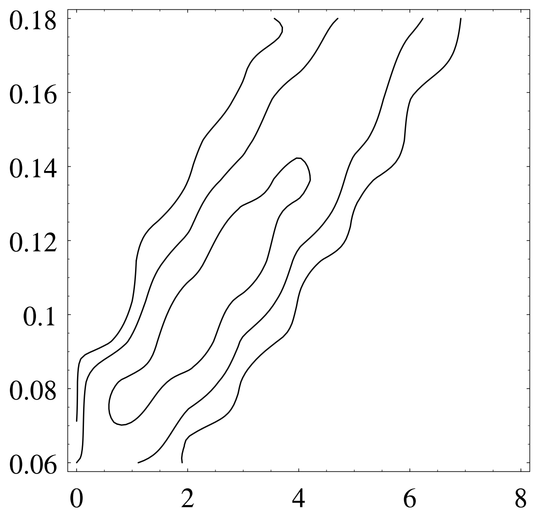

Fig. 1. Likelihood contours from WMAP data in the N_{eff} (x-axis) - $\Omega_{cdm}h^2$ (y-axis) plane [5].

3 $\Omega_b h^2$ and N_{eff} after WMAP

The peak structure of the CMB power spectrum has been beautifully confirmed by a series of experiment in the past few years (BOOMERanG, MAXIMA, DASI, CBI, ACBAR) and more recently by WMAP collaboration [2] with a very high accuracy. The improvement in angular resolution from the 7 degrees across the sky of COBE to the order 0.5 degrees of WMAP, allows us to have a better understanding of several features of our Universe and in particular of its matter-energy content in terms of cosmological constant, dark matter and baryons.

The role of relativistic species at the CMB formation, at redshifts of the order of $z \sim 1100$, is mainly to shift the matter radiation equality time, which results in both shifting the peak location in angular scale and changing the power around the first acoustic peak. This is basically due to a change in the early integrated Sachs-Wolfe effect. Several groups have studied this topics, obtaining comparable bounds on N_{eff} but using different priors [3]-[5]. For example, in our analysis [5], $N_{eff} = 2.6^{+3.5}_{-2.0}$, using WMAP data only and weak prior on the value of the Hubble parameter, $h = 0.7 \pm 0.2$. The reason for such a wide range for N_{eff} is ultimately due to the many unknown cosmological parameters which determine the power spectrum, and in particular to the presence of several degeneracies, i.e. the fact that different choices for some of these parameters produce the very same power spectrum. As an example if we increase both N_{eff} and the amount of dark matter Ω_{cdm} we can obtain the same power spectrum provided we do not change the radiation-matter equality. This is shown in Fig.1, a plot of the bi-dimensional likelihood contours in the $N_{eff} - \Omega_{cdm}$ plane

The baryon density parameter $\Omega_b h^2$ can be much more severely constrained from the power spectrum. Increasing the baryons in the plasma enhances the ef-

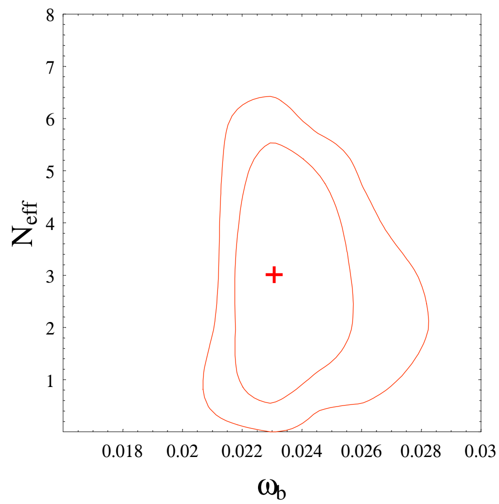

Fig. 2. The 68 and 95% C.L. likelihood contours from WMAP data in the $N_{eff} - \omega_b$ plane.

fective mass of the fluid and this leads to more pronounced compression peaks. By a likelihood analysis the bound obtained in [5] is $\Omega_b h^2 = 0.023 \pm 0.002$ $(1-\sigma$ error), fully compatible with the one quoted by the WMAP Collaboration, $\Omega_b h^2 = 0.022 \pm 0.001$, [2]. In Fig.2 we show the likelihood contours in the $N_{eff} - \Omega_b h^2$ plane. This result is extremely important. The WMAP data tell us the value of baryon density with a better accuracy than BBN, so we can test the standard scenario of light nuclei formation with basically no free parameters but the value of N_{eff}.

4 $\Omega_b h^2$ and N_{eff} and Big Bang Nuclesynthesis

The primordial production of light nuclei, mainly ^4He, D and ^7Li, takes place when the temperature of the electromagnetic plasma is in the range $1 \div 0.01 MeV$, and is strongly influenced by the two parameters $\Omega_b h^2$ and N_{eff}. Increasing the value of the baryon to photon number density enhances the fusion mechanism, so it leads to a larger eventual amount of ^4He, the most tightly bound light nucleus (^4He binding energy per nucleon is of the order of $7MeV$). On the other hand, D rapidly decreases with $\Omega_b h^2$, so the experimental result on this species is a very sensitive measure of baryons in the universe. The contribution of relativistic degrees of freedom to the expansion rate, parameterized by N_{eff} affects instead the decoupling temperature of weak reaction which keep in chemical equilibrium protons and neutrons. For large temperatures in fact the ratio of their densities is given by equilibrium conditions, $n/p = \exp(-(m_n - m_p)/T)$, therefore if weak interactions were efficient down to very low temperatures, much smaller than the neutron-proton mass difference, neutrons would completely disappear. We mentioned however that the rate of these processes indeed becomes smaller than the expansion rate H for temperatures of the order of $T_D \sim 1MeV$, so that the

n/p ratio freezes-out at the value $n/p = \exp(-(m_n - m_p)/T_D)$. Since almost all neutrons are eventually bound in ^4He nuclei it is then straightforward to get for the Helium mass fraction

$$Y_p \equiv \frac{4n_{He}}{n_b} \sim 2\frac{n/p}{1+n/p} = 2\frac{1}{e^{(m_n-m_p)/T_D}+1} \tag{1}$$

When correcting this result for neutron spontaneous decay one gets $Y_p \sim 0.25$, already an excellent estimate compared with the result of detailed numerical calculations. Changing N_{eff} affect the decoupling temperature T_D and so the amount of primordial Helium.

An accurate analysis of BBN can be only achieved by numerically solving a set of coupled differential equations, taking into account quite a complicated network of nuclear reactions. Some of these reactions are still affected by large uncertainties, which therefore introduce an error in the theoretical prediction for, mainly, D and ^7Li abundances. As we mentioned Helium prediction is mainly influenced by $n \leftrightarrow p$ processes, which are presently known at a high level of accuracy [6]-[7]. Quite recently a big effort has been devoted in trying to quantify the role of each nuclear reaction to the uncertainties on nuclei abundances, using either Monte Carlo [8] or linear propagation [9] techniques. The most recent analysis [5], [10], [11] have benefited from the NACRE nuclear reaction catalogue [12], as well as of very recent results, as for example the LUNA Collaboration measurement of the $D(p,\gamma)^3He$ [13]. We report here the results obtained in [5] for the total relative theoretical uncertainties σ_i^{th} on Y_p and D and ^7Li number fractions $X_i = n_i/n_b$

$$\frac{\sigma_D}{X_D} \sim 10\%, \quad \frac{\sigma_{He}}{Y_p} \sim 0.1\%, \quad \frac{\sigma_{Li}}{X_{Li}} \sim 25\% \tag{2}$$

The large error on ^7Li is mainly due to the uncertainty on the rate for the process $^4He(^3He,\gamma)^7Be$, a process which is also of great interest for the determination of both ^7Be and ^8B neutrino fluxes from the sun. Hopefully it will be studied at low energies in the near future.

The experimental determination of primordial abundances is really a challenging task. The strategy is to identify metal poor environment, which are not been severely polluted by star contamination in their light nuclei content, and possibly to correct the observations for the effect of galactic evolution.

The ^4He mass fraction is determined by regression to zero metallicity of the values obtained by observing $HeII \rightarrow HeI$ recombination lines in extragalactic ionized gas. There are still quite different results (see e.g. [14] for a review and references), a low one, $Y_p = 0.234 \pm 0.003$, and a $high$ value, $Y_p = 0.244 \pm 0.002$. In the following we also use a conservative estimate, $Y_p = 0.239 \pm 0.008$.

The best estimate of Primordial D comes from observations of absorption lines in gas clouds in the line of sight between the earth and Quasars at very high redshift ($z \sim 2-3$), which give $X_D = (2.78^{+0.44}_{-0.38}) \cdot 10^{-5}$ [15]. Finally ^7Li is measured via observation of absorption lines in spectra of POP II halo stars, which show a saturation of ^7Li abundance at low metallicity (Spite plateau).

The present status of BBN, in the standard scenario, using the value of baryon density as determined by WMAP and $N_{eff} = 3.04$ is summarized in Fig.s 3-5.

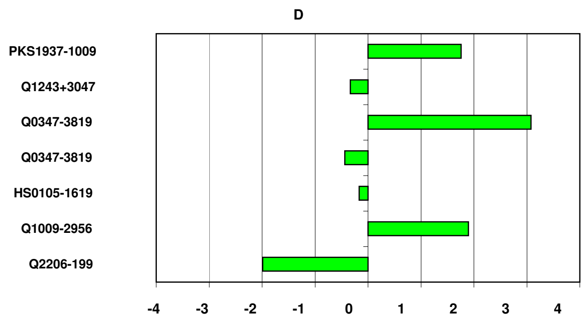

Fig. 3. The pulls of QSO D measurements with respect to the theoretical prediction for $N_{eff} = 3.04$ and $\omega_b = 0.023$, in units of $\left((\sigma_{DD}^{th})^2 + (\sigma_D^{exp})^2\right)^{1/2}$.

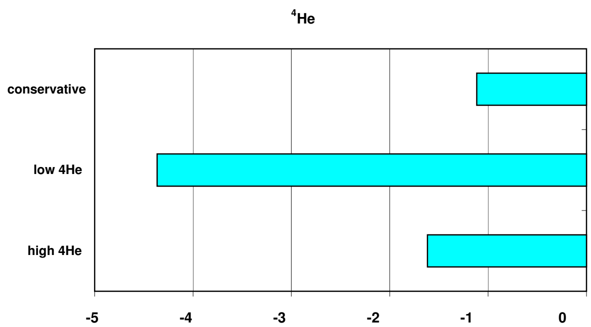

Fig. 4. The pulls of Y_p measurements with respect to the theoretical prediction for $N_{eff} = 3.04$ and $\omega_b = 0.023$, in units of $\left((\sigma_{44}^{th})^2 + (\sigma_4^{exp})^2\right)^{1/2}$.

Here I report the difference between the theoretical and the experimental determination, normalized to the total uncertainty, theoretical and experimental, summed in quadrature. The average of the several D measurements, reported above, is indeed in very good agreement with theory. This is a very crucial result since, as we said already, D is strongly influenced by $\Omega_b h^2$, which is now fixed by WMAP. In Fig. 6 I show the combined likelihood contours at 2σ in the $N_{eff} - \Omega_b h^2$ plane obtained when using the WMAP result and D measurement only (colored area) and the D $+^4$ He results using the conservative Y_p shown before.

It is evident that the effect of ^4He is to shift the values of both $\Omega_b h^2$ and N_{eff} towards smaller values, which produces a smaller theoretical value for Y_p. Though this may be seen as a (weak) indication of the fact that a slightly lower value for N_{eff} is preferred, I would more conservatively say that, waiting for a more clear understanding of possible systematics in Y_p experimental determination, the standard scenario for BBN is in reasonable good shape. An open problem is however still represented by the evidence for ^7Li depletion, which is not fully

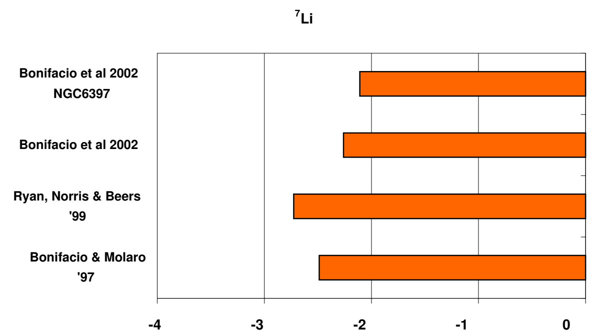

Fig. 5. The pulls of X_{7Li} measurements with respect to the theoretical prediction for $N_{eff} = 3.04$ and $\omega_b = 0.023$, in units of $\left((\sigma^{th}_{77})^2 + (\sigma^{exp}_7)^2\right)^{1/2}$.

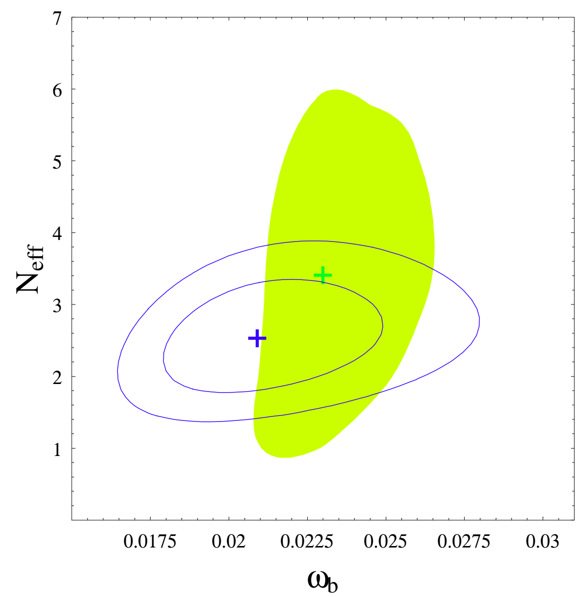

Fig. 6. The 68 and 95% C.L. contours for the $D +^4 He$ likelihood function in the $\omega_b - N_{eff}$ plane ($\omega_b = \Omega_b h^2$). We also show the result of the $CMB + D$ analysis (colored area).

understood (see Fig. 5). The theoretical result for X_{Li} is in fact a factor 2-3 larger than the present experimental determination.

5 Neutrino-antineutrino asymmetry

While the electron-positron asymmetry density is severely constrained, of the order of 10^{-10} in unit of the photon density, we have no bounds at all on neutrino asymmetry from charge neutrality of the universe. Defining $\xi_x = \mu_x/T_x$, with μ_x the chemical potential for the ν_x species, with $x = e, \mu, \tau$, we recall that

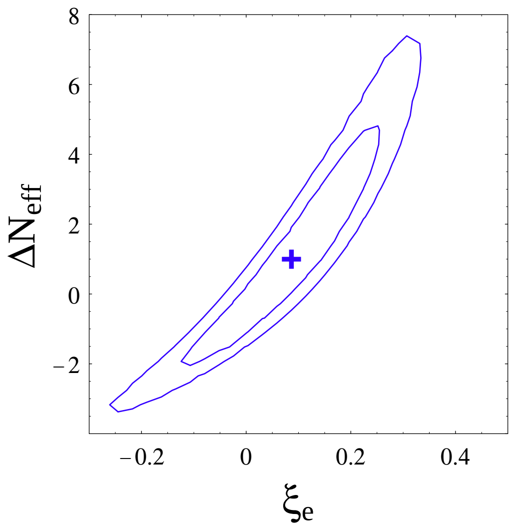

Fig. 7. The 68 and 95% C.L. likelihood contours in the $\xi_e - \Delta N_{eff}$ plane from BBN, with $\Delta N_{eff} = N_{eff} - 3.04$.

for a Fermi-Dirac distribution the particle-antiparticle asymmetry is simply related to ξ_x (I assume here for simplicity massless neutrinos)

$$n(\nu_x) - n(\overline{\nu}_x) = \frac{T_x^3}{6}\left(\xi_x + \frac{\xi_x^3}{\pi^2}\right) \tag{1}$$

since neutrinos decoupled as hot relics starting from a chemical equilibrium condition with $e^{\pm}$, so that $\mu_x \equiv \mu(\nu_x) = -\mu(\overline{\nu}_x)$. Non vanishing values for ξ_x affects very weakly CMB, while it is much more constrained by BBN. In fact any asymmetry in the neutrino sector contribute to the Hubble expansion rate, i.e. to N_{eff}

$$N_{eff} \to N_{eff} + \sum_x \left[\frac{30}{7}\left(\frac{\xi_x}{\pi}\right)^2 + \frac{15}{7}\left(\frac{\xi_x}{\pi}\right)^4\right] \tag{2}$$

In addition the asymmetry in the electron neutrino sector directly affects the n/p value at the freeze-out of weak interactions, since they directly enter in the processes governing this phenomenon, namely $n + \nu_e \leftrightarrow p + e^-$, $n \leftrightarrow p + e^- + \overline{\nu}_e$ and $n + e^+ \leftrightarrow p + \overline{\nu}_e$.

It was recently realized [16] that indeed, because of flavor oscillation, using present determination of mass differences and mixing angles from atmospheric and solar neutrinos, the three ξ_x should be very close each other, so the bound on their (common) value ξ mainly come from the fact that ξ_e should be quite small $\xi \leq 0.1$, to give a value for the n/p ratio (and so for ^{4}He) in agreement with data. In Fig. 7 I show the likelihood contour obtained in the $\xi - N_{eff}$ plane [5]. Though the standard BBN is preferred, there is still room for very exotic scenarios, with larger neutrino degeneracies and even very large (or very small) N_{eff}.

6 Cosmological bounds on neutrino mass scale

Despite of the fact that we presently know neutrino mass differences from oscillation effects in atmospheric and solar neutrino fluxes, there is still quite a wide range for their absolute mass scale, spanning several order of magnitude, from few eV down to 10^{-2} eV. Terrestrial bounds come from Tritium decay experiments [17], which presently give $m(\nu_e) \leq 2.2eV$. This result will be greatly improved by next generation experiment KATRIN, which should reach a sensitivity after three years of running of the order of 0.35 eV [18]

An independent source of information will be provided by neutrinoless beta decay, which is sensitive to the effective ν_e mass

$$< m_e >= |U_{e1}|^2 m_1 + |U_{e2}|^2 e^{i\phi_2} m_2 + |U_{e3}|^2 e^{i\phi_3} m_3 \tag{1}$$

with U_{ei} the electron neutrino projection onto mass eigenstates with mass m_i, and ϕ_i CP violating Majorana phases. Planned experiments CUORE [19] and GENIUS [20] will have a sensitivity on this parameter of the order of $10^{-1} - 10^{-2}$ eV.

Interestingly, quite severe constraints on neutrino masses come from cosmology. Massive neutrinos in fact contribute to the present total energy density of the Universe as $m_\nu n_\nu$ so we get

$$\Omega_\nu h^2 = \frac{\sum_x m(\nu_x)}{92.5\,\mathrm{eV}} \tag{2}$$

which gives a generous bound when imposing $\Omega_\nu h^2 < 1$.

Neutrino masses however also enter in the way structures grow for gravitational instability from the initial seed likely given by adiabatic perturbations produced during the inflationary phase. In fact neutrinos free stream and tend to suppress structure formation on all scales unless they are massive. In this case they can only affect scales smaller than the Hubble horizon when they eventually become non relativistic, the ones with a corresponding wave number larger than [21]

$$k_{nr} \sim 0.026 \left(\frac{m_\nu}{1eV}\right)^{1/2} \Omega_m^{1/2} h \mathrm{Mpc}^{-1} \tag{3}$$

Several authors have recently considered this issue in details [22]- [24], combining WMAP data and the results of the 2dFGRS survey [25]. Actually the result also depends on the specific value of N_{eff}. A conservative value is given by $\sum_x m(\nu_x) \leq 2\,\mathrm{eV}$.

7 Acknowledgments

It is a pleasure to thank Norma Mankoc Borstnik for organizing the Portoroz meeting and for her warm hospitality. I also thank my collaborators at Naples University A. Cuoco, F. Iocco, G. Miele, O. Pisanti and P. Serpico.

References

1. G. Mangano, G. Miele, S. Pastor, and M. Peloso, Phys. Lett. B534, 8 (2002).
2. C.L. Bennett *et al.* (WMAP Coll.), Astrophys. J. Suppl. 148, 1 (2003).
3. P. Crotty, J. Lesgourgues, and S. Pastor, Phys. Rev. D67, 123005 (2003).
4. E. Pierpaoli, Mon. Not. Roy. Astron. Soc. 342, L63 (2003).
5. A. Cuoco *et al.* ., astro-ph/0307213.
6. R.E. Lopez and M.S. Turner, Phys. Rev. D59, 103502 (1999).
7. S. Esposito, G. Mangano, G. Miele and O. Pisanti, Nucl. Phys. B568, 421 (2000).
8. L.M. Krauss and P. Romanelli, Astrophys. J. 358, 47 (1990); P.J. Kernan and L.M. Krauss, Phys. Rev. Lett. 72, 3309 (1994).
9. G. Fiorentini, E. Lisi, S. Sarkar, and F.L. Villante, Phys. Rev. D58, 063506 (1998).
10. S. Burles, K.M. Nollett, and M.S. Turner, Astrophys. J. 552, L1 (2001).
11. A. Coc, E. Vangioni-Flam, and M. Cassé, astro-ph/0002248; astro-ph/0009297; astro-ph/0101286.
12. C. Angulo *et al.* , *Nucl. Phys.* **A656** (1999) 3. **NACRE** web site *http://pntpm.ulb.ac.be/nacre.htm*
13. LUNA Collaboration, *Nucl. Phys.* **A706** (2002) 203.
14. B.D. Fields, and S. Sarkar, in *Review of Particle Properties*, Phys. Rev. D66, 010001 (2002).
15. D. Kirkman, D. Tytler, N. Suzuki, J.M. O'Meara, and D. Lubin, astro-ph/0302006.
16. A.D. Dolgov, S.H. Hansen, S. Pastor, S.T. Petcov, G.G. Raffelt, and D.V. Semikoz, Nucl. Phys. B362, 363 (2002).
17. V.M. Lobashev *et al.* (Troisk Coll.), Phys. Lett. B460, 227 (1999); C. Weinheimer *et al.* (Mainz Coll.), Phys. Lett. B460, 219 (1999).
18. A. Osipowicz *et al.* , hep-ph/0109033.
19. C. Arnaboldi *et al.* (CUORE Coll.), hep-ex/0212053.
20. H.V. Klapdor-Kleingrothaus (GENIUS Coll.), hep-ph/9910205.
21. W. Hu, D.J. Eisenstein, and M. Tegmark, Phys. Rev. Lett. 80, 5255 (1998).
22. D.N. Spergel *et al.* , Astrophys. J. Suppl. 148, 175 (2003).
23. O. Elgaroy and O. Lahav, JCAP 0304, 004 (2003).
24. S. Hannestad, JCAP 0305, 004 (2003).
25. J. Peacock *et al.* , Nature 410, 169 (2001).

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 171)*
Portorož, Slovenia, July 12-17, 2003

The Problem of Mass

Colin D. Froggatt*

Department of Physics and Astronomy, Glasgow University,
Glasgow G12 8QQ, Scotland, UK

Abstract. The quark-lepton mass problem and the ideas of mass protection are reviewed. The Multiple Point Principle is introduced and used within the Standard Model to predict the top quark and Higgs particle masses. We discuss the lightest family mass generation model, in which all the quark mixing angles are successfully expressed in terms of simple expressions involving quark mass ratios. The chiral flavour symmetry of the family replicated gauge group model is shown to provide the mass protection needed to generate the hierarchical structure of the quark-lepton mass matrices.

1 Introduction

The most important unresolved problem in particle physics is the understanding of flavour and the fermion mass spectrum. The observed values of the fermion masses and mixing angles constitute the bulk of the Standard Model (SM) parameters and provide our main experimental clues to the underlying flavour dynamics. In particular the non-vanishing neutrino masses and mixings provide direct evidence for physics beyond the SM.

The charged lepton masses can be directly measured and correspond to the poles in their propagators:

$$M_e = 0.511 \text{ MeV} \qquad M_\mu = 106 \text{ MeV} \qquad M_\tau = 1.78 \text{ GeV} \tag{1}$$

However the quark masses have to be extracted from the properties of hadrons and are usually quoted as running masses $m_q(\mu)$ evaluated at some renormalisation scale μ, which are related to the propagator pole masses M_q by

$$M_q = m_q(\mu = m_q)\left[1 + \frac{4}{3}\alpha_3(m_q)\right] \tag{2}$$

to leading order in QCD. The light u, d and s quark masses are usually normalised to the scale $\mu = 1$ GeV (or $\mu = 2$ GeV for lattice measurements) and to the quark mass itself for the heavy c, b and t quarks. They are typically given

* E-mail: c.froggatt@physics.gla.ac.uk

[1] as follows[1]:

$$\begin{aligned} m_u(1\text{ GeV}) &= 4.5 \pm 1\text{ MeV} & m_d(1\text{ GeV}) &= 8 \pm 2\text{ MeV} \\ m_c(m_c) &= 1.25 \pm 0.15\text{ GeV} & m_s(1\text{ GeV}) &= 150 \pm 50\text{ MeV} \\ m_t(m_t) &= 166 \pm 5\text{ GeV} & m_b(m_b) &= 4.25 \pm 0.15\text{ GeV} \end{aligned} \tag{3}$$

However we only have an upper limit on the neutrino masses of $m_{\nu_i} \lesssim 1$ eV from tritium beta decay and from cosmology, and measurements of the neutrino mass squared differences:

$$\Delta m_{21}^2 \sim 5 \times 10^{-5}\text{eV}^2 \qquad \Delta m_{32}^2 \sim 3 \times 10^{-3}\text{eV}^2 \tag{4}$$

from solar and atmospheric neutrino oscillation data [2].

The magnitudes of the quark mixing matrix V_{CKM} are well measured

$$|V_{CKM}| = \begin{pmatrix} 0.9734 \pm 0.0008 & 0.2196 \pm 0.0020 & 0.0036 \pm 0.0007 \\ 0.224 \pm 0.016 & 0.996 \pm 0.013 & 0.0412 \pm 0.002 \\ 0.0077 \pm 0.0014 & 0.0397 \pm 0.0033 & 0.9992 \pm 0.0002 \end{pmatrix} \tag{5}$$

and a CP violating phase of order unity:

$$\sin^2 \delta_{CP} \sim 1 \tag{6}$$

can reproduce all the CP violation data. Neutrino oscillation data constrain the magnitudes of the lepton mixing matrix elements to lie in the following 3σ ranges [2]:

$$|U_{MNS}| = \begin{pmatrix} 0.73 - 0.89 & 0.45 - 0.66 & < 0.24 \\ 0.23 - 0.66 & 0.24 - 0.75 & 0.52 - 0.87 \\ 0.06 - 0.57 & 0.40 - 0.82 & 0.48 - 0.85 \end{pmatrix} \tag{7}$$

Due to the Majorana nature of the neutrino mass matrix, there are three unknown CP violating phases δ, α_1 and α_2 in this case [2].

The charged fermion masses range over five orders of magnitude, whereas there seems to be a relatively mild neutrino mass hierarchy. The absolute neutrino mass scale ($m_\nu < 1$ eV) suggests a new physics mass scale – the so-called see-saw scale $\Lambda_{seesaw} \sim 10^{15}$ GeV. The quark mixing matrix V_{CKM} is also hierarchical, with small off-diagonal elements. However the elements of U_{MNS} are all of the same order of magnitude except for $|U_{e3}| < 0.24$, corresponding to two leptonic mixing angles being close to maximal ($\theta_{atmospheric} \simeq \pi/4$ and $\theta_{solar} \simeq \pi/6$).

We introduce the mechanism of mass protection by approximately conserved chiral charges in section 2. The top quark mass is the dominant term in the SM fermion mass matrix, so it is likely that its value will be understood dynamically before those of the other fermions. In section 3 we discuss the connection between the top quark and Higgs masses and how they can be determined from the so-called Multiple Point Principle. We present the lightest family mass generation model in section 4, which provides an ansatz for the texture of fermion mass

[1] Note that the top quark mass, $M_t = 174 \pm 5$ GeV, measured at FermiLab is interpreted as the pole mass.

matrices and expresses all the quark mixing angles successfully in terms of simple expressions involving quark mass ratios. The family replicated gauge group model is presented in section 5, as an example of a model whose gauge group naturally provides the mass protecting quantum numbers needed to generate the required texture for the fermion mass matrices. Finally we present a brief conclusion in section 6.

2 Fermion Mass and Mass Protection

A fermion mass term

$$\mathcal{L}_{mass} = -m\overline{\psi}_L\psi_R + h.c. \tag{8}$$

couples together a left-handed Weyl field ψ_L and a right-handed Weyl field ψ_R, which then satisfy the Dirac equation

$$i\gamma^\mu\partial_\mu\psi_L = m\psi_R \tag{9}$$

If the two Weyl fields are not charge conjugates $\psi_L \neq (\psi_R)^c$ we have a Dirac mass term and the two fields ψ_L and ψ_R together correspond to a Dirac spinor. However if the two Weyl fields are charge conjugates $\psi_L = (\psi_R)^c$ we have a Majorana mass term and the corresponding four component Majorana spinor has only two degrees of freedom. Particles carrying an exactly conserved charge, like the electron carries electric charge, must be distinct from their anti-particles and can only have Dirac masses with ψ_L and ψ_R having equal conserved charges. However a neutrino could be a Majorana particle.

If ψ_L and ψ_R have different quantum numbers, i.e. belong to inequivalent representations of a symmetry group G (G is then called a chiral symmetry), a Dirac mass term is forbidden in the limit of an exact G symmetry and they represent two massless Weyl particles. Thus the G symmetry “protects” the fermion from gaining a mass. Such a fermion can only gain a mass when G is spontaneously broken.

The left-handed and right-handed top quark, t_L and t_R, carry unequal Standard Model $SU(2) \times U(1)$ gauge charges $\mathbf{Q}$:

$$\mathbf{Q}_L \neq \mathbf{Q}_R \qquad \text{(Chiral charges)} \tag{10}$$

Hence electroweak gauge invariance protects the quarks and leptons from gaining a fundamental mass term ($\bar{t}_L t_R$ is not gauge invariant). This *mass protection* mechanism is of course broken by the Higgs effect, when the vacuum expectation value of the Weinberg-Salam Higgs field

$$< \phi_{WS} >= \sqrt{2}\nu = 246\ \text{GeV} \tag{11}$$

breaks the gauge symmetry and the SM gauge invariant Yukawa couplings $\frac{y_i}{\sqrt{2}}$ generate the running quark masses $m_i = y_i\nu = 174\, y_i$ GeV. In this way a top quark mass of the same order of magnitude as the SM Higgs field vacuum expectation value (VEV) is naturally generated (with y_t unsuppressed). Thus the Higgs mechanism explains why the top quark mass is suppressed, relative to the

fundamental (Planck, GUT...) mass scale of the physics beyond the SM, down to the scale of electroweak gauge symmetry breaking. However the further suppression of the other quark-lepton masses ($y_b, y_c, y_s, y_u, y_d \ll 1$) remains a mystery, which it is natural to attribute to mass protection by other approximately conserved chiral gauge charges beyond the SM, as discussed in section 5 for the family replicated gauge group model.

Fermions which are vector-like under the SM gauge group ($\mathbf{Q}_L = \mathbf{Q}_R$) are not mass protected and are expected to have a large mass associated with new (grand unified, string,..) physics. The Higgs particle, being a scalar, is not mass protected and *a priori* would also be expected to have a large mass; this is the well-known gauge hierarchy problem discussed at Portoroz by Holger Nielsen [3].

3 Top Quark and Higgs Masses from the Multiple Point Principle

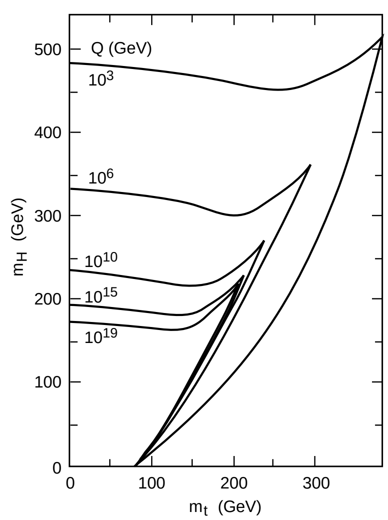

Fig. 1. SM bounds in the (m_t,m_H) plane for various values of $\Lambda = Q$, the scale at which new physics enters.

It is well-known [4] that the self-consistency of the pure SM up to some physical cut-off scale Λ imposes constraints on the top quark and Higgs boson masses. The first constraint is the so-called triviality bound: the running Higgs coupling constant $\lambda(\mu)$ should not develop a Landau pole for $\mu < \Lambda$. The second is the vacuum stability bound: the running Higgs coupling constant $\lambda(\mu)$ should not become negative leading to the instability of the usual SM vacuum. These bounds are illustrated [5] in Fig. 1, where the combined triviality and vacuum stability bounds for the SM are shown for different values of the high energy cut-off Λ. The allowed region is the area around the origin bounded by the co-ordinate axes

and the solid curve labelled by the appropriate value of Λ. The upper part of each curve corresponds to the triviality bound. The lower part of each curve coincides with the vacuum stability bound and the point in the top right hand corner, where it meets the triviality bound curve, is the infra-red quasi-fixed point for that value of Λ. Here the vacuum stability curve, for a large cut-off of order the Planck scale $\Lambda_{\mathtt{Planck}} \simeq 10^{19}$ GeV, is important for the discussion of the values of the top quark and Higgs boson masses predicted from the Multiple Point Principle.

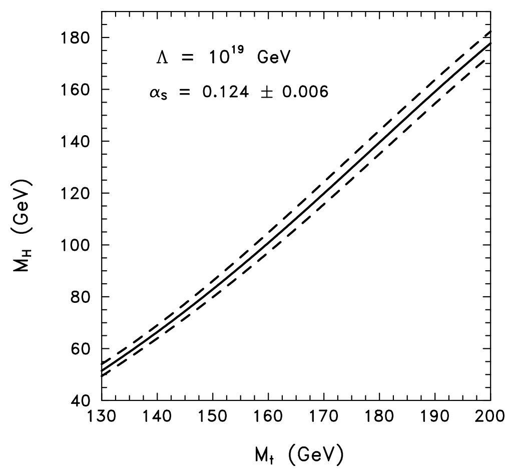

Fig. 2. SM vacuum stability curve for $\Lambda = 10^{19}$ GeV and $\alpha_s = 0.124$ (solid line), $\alpha_s = 0.118$ (upper dashed line), $\alpha_s = 0.130$ (lower dashed line).

According to the Multiple Point Principle (MPP), Nature chooses coupling constant values such that a number of vacuum states have the same energy density (cosmological constant). This fine-tuning of the coupling constants is similar to that of temperature for a mixture of co-existing phases such as ice and water. We have previously argued [6] that baby-universe like theories [7], having a mild breaking of locality and causality, may contain the underlying physical explanation of the MPP, but it really has the status of a postulated new principle. Here we apply it to the pure Standard Model [8], which we assume valid up close to $\Lambda_{\mathtt{Planck}}$. So we shall postulate that the effective potential $V_{eff}(\phi)$ for the SM Higgs field ϕ should have a second minimum, at $<\phi>=\phi_{vac\ 2}$, degenerate with the well-known first minimum at the electroweak scale $<\phi>=\phi_{vac\ 1} = 246$ GeV:

$$V_{eff}(\phi_{vac\ 1}) = V_{eff}(\phi_{vac\ 2}) \tag{12}$$

Thus we predict that our vacuum is barely stable and we just lie on the vacuum stability curve in the top quark, Higgs particle (pole) mass (M_t, M_H) plane, shown [9] in Fig. 2 for a cut-off $\Lambda = 10^{19}$ GeV. Furthermore we expect the second minimum to be within an order of magnitude or so of the fundamental scale, i.e. $\phi_{vac\ 2} \simeq \Lambda_{\mathtt{Planck}}$. In this way, we essentially select a particular point on the SM

vacuum stability curve and hence the MPP condition predicts precise values for M_t and M_H.

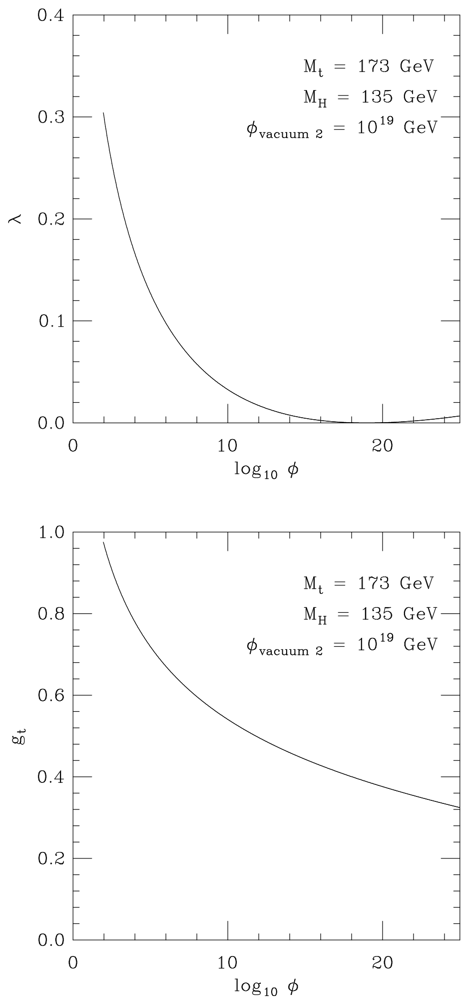

Fig. 3. Plots of λ and g_t as functions of the scale of the Higgs field ϕ for degenerate vacua with the second Higgs VEV at the Planck scale $\phi_{vac\ 2} = 10^{19}$ GeV. We formally apply the second order SM renormalisation group equations up to a scale of 10^{25} GeV.

For large values of the SM Higgs field $\phi >> \phi_{vac\ 1}$, the renormalisation group improved tree level effective potential is very well approximated by

$$V_{eff}(\phi) \simeq \frac{1}{8}\lambda(\mu = |\phi|)|\phi|^4$$

and the degeneracy condition, eq. (12), means that $\lambda(\phi_{vac\ 2})$ should vanish to high accuracy. The derivative of the effective potential $V_{eff}(\phi)$ should also be

zero at $\phi_{vac\ 2}$, because it has a minimum there. Thus at the second minimum of the effective potential the beta function β_λ vanishes as well:

$$\beta_\lambda(\mu = \phi_{vac\ 2}) = \lambda(\phi_{vac\ 2}) = 0 \tag{13}$$

which gives to leading order the relationship:

$$\frac{9}{4}g_2^4 + \frac{3}{2}g_2^2 g_1^2 + \frac{3}{4}g_1^4 - 12g_t^4 = 0 \tag{14}$$

between the top quark Yukawa coupling $g_t(\mu)$ and the electroweak gauge coupling constants $g_1(\mu)$ and $g_2(\mu)$ at the scale $\mu = \phi_{vac\ 2} \simeq \Lambda_{Planck}$. We use the renormalisation group equations to relate the couplings at the Planck scale to their values at the electroweak scale. Figure 3 shows the running coupling constants $\lambda(\phi)$ and $g_t(\phi)$ as functions of $\log(\phi)$. Their values at the electroweak scale give our predicted combination of pole masses [8]:

$$M_t = 173 \pm 5 \text{ GeV} \quad M_H = 135 \pm 9 \text{ GeV} \tag{15}$$

We have also considered [10] a slightly modified version of MPP, according to which the two vacua are approximately degenerate in such a way that they should both be physically realised over comparable amounts of space-time four volume. This modified MPP corresponds to the Higgs mass lying on the vacuum *metastability* curve rather than on the vacuum stability curve, giving a Higgs mass prediction of 122 ± 11 GeV. We should presumably not really take the MPP predictions to be more accurate than to the order of magnitude of the variation between the metastability and stability bounds. However we definitely predict a light Higgs mass in this range, as seems to be in agreement with indirect estimates of the SM Higgs mass from precision data [1].

This application of the MPP assumes the existence of the hierarchy

$$v/\Lambda_{Planck} \sim 10^{-17}.$$

Recently we have speculated [11] that this huge scale ratio is a consequence of the existence of yet another vacuum in the SM, at the electroweak scale and degenerate with the two vacua discussed above. The two SM vacua at the electroweak scale are postulated to differ by the condensation of an S-wave bound state formed from 6 top and 6 anti-top quarks mainly due to Higgs boson exchange forces. This scenario is discussed in more detail in Holger Nielsen's talk [3].

4 Lightest Family Mass Generation Model

Motivated by the famous Fritzsch ansatz [12] for the two generation quark mass matrices:

$$M_U = \begin{pmatrix} 0 & B \\ B^* & A \end{pmatrix} \qquad M_D = \begin{pmatrix} 0 & B' \\ B'^* & A' \end{pmatrix} \tag{16}$$

several ansätze have been proposed for the fermion mass matrices—for example, see [13] for a systematic analysis of symmetric quark mass matrices with texture

zeros at the SUSY-GUT scale. Here I will concentrate on the lightest family mass generation model [14]. It successfully generalizes the well-known formula

$$|V_{us}| \simeq \left| \sqrt{\frac{m_d}{m_s}} - e^{i\phi} \sqrt{\frac{m_u}{m_c}} \right| \tag{17}$$

for the Cabibbo angle derived from the above ansatz, eq. (16), to simple working formulae for all the quark mixing angles in terms of quark mass ratios. According to this model the flavour mixing for quarks is basically determined by the mechanism responsible for generating the physical masses of the up and down quarks, m_u and m_d respectively. So, in the chiral symmetry limit, when m_u and m_d vanish, all the quark mixing angles vanish. Therefore we are led to consider an ansatz in which the diagonal mass matrix elements for the second and third generations are practically the same in the gauge (unrotated) and physical bases.

The mass matrix for the down quarks ($D = d, s, b$) is taken to be hermitian with three texture zeros of the following form:

$$M_D = \begin{pmatrix} 0 & a_D & 0 \\ a_D^* & A_D & b_D \\ 0 & b_D^* & B_D \end{pmatrix} \tag{18}$$

where

$$B_D = m_b + \delta_D \qquad A_D = m_s + \delta'_D \qquad |\delta_D| \ll m_s \qquad |\delta'_D| \ll m_d \tag{19}$$

It is, of course, necessary to assume some hierarchy between the elements, which we take to be: $B_D \gg A_D \sim |b_D| \gg |a_D|$. The zero in the $(M_D)_{11}$ element corresponds to the commonly accepted conjecture that the lightest family masses appear as a direct result of flavour mixings. The zero in $(M_D)_{13}$ means that only minimal "nearest neighbour" interactions occur, giving a tridiagonal matrix structure. Since the trace and determinant of the hermitian matrix M_D gives the sum and product of its eigenvalues, it follows that

$$\delta_D \simeq -m_d \tag{20}$$

while δ'_D is vanishingly small and can be neglected in further considerations.

It may easily be shown that equations (18 - 20) are entirely equivalent to the condition that the diagonal elements (A_D, B_D) of M_D are proportional to the modulus square of the off-diagonal elements (a_D, b_D):

$$\frac{A_D}{B_D} = \left| \frac{a_D}{b_D} \right|^2 \tag{21}$$

Using the conservation of the trace, determinant and sum of principal minors of the hermitian matrix M_D under unitary transformations, we are led to a complete determination of the moduli of all its elements, which can be expressed to high accuracy as follows:

$$|M_D| = \begin{pmatrix} 0 & \sqrt{m_d m_s} & 0 \\ \sqrt{m_d m_s} & m_s & \sqrt{m_d m_b} \\ 0 & \sqrt{m_d m_b} & m_b - m_d \end{pmatrix} \tag{22}$$

The mass matrix for the up quarks is taken to be of the following hermitian form:

$$M_U = \begin{pmatrix} 0 & 0 & c_U \\ 0 & A_U & 0 \\ c_U^* & 0 & B_U \end{pmatrix} \tag{23}$$

The moduli of all the elements of M_U can also be readily determined in terms of the physical masses as follows:

$$|M_U| = \begin{pmatrix} 0 & 0 & \sqrt{m_u m_t} \\ 0 & m_c & 0 \\ \sqrt{m_u m_t} & 0 & m_t - m_u \end{pmatrix} \tag{24}$$

The CKM quark mixing matrix elements can now be readily calculated by diagonalising the mass matrices M_D and M_U. They are given in terms of quark mass ratios as follows:

$$|V_{us}| = \sqrt{\frac{m_d}{m_s}} = 0.222 \pm 0.004 \qquad |V_{us}|_{exp} = 0.221 \pm 0.003 \tag{25}$$

$$|V_{cb}| = \sqrt{\frac{m_d}{m_b}} = 0.038 \pm 0.004 \qquad |V_{cb}|_{exp} = 0.039 \pm 0.003 \tag{26}$$

$$|V_{ub}| = \sqrt{\frac{m_u}{m_t}} = 0.0036 \pm 0.0006 \qquad |V_{ub}|_{exp} = 0.0036 \pm 0.0006 \tag{27}$$

$$|V_{td}| = |V_{us}V_{cb} - V_{ub}| = 0.009 \pm 0.002 \qquad |V_{ub}|_{exp} = 0.0077 \pm 0.0014 \tag{28}$$

As can be seen, they are in impressive agreement with the experimental values. The MNS lepton mixing matrix can also be fitted, if the texture of eq. (18) is extended to the Dirac and Majorana right-handed neutrino mass matrices [15].

The proportionality condition, eq. (21), is not so easy to generate from an underlying symmetry beyond the Standard Model, but it is possible to realise it in a local chiral $SU(3)$ family symmetry[2] model [16].

5 Family Replicated Gauge Group Model

As pointed out in section 2, a natural explanation of the charged fermion mass hierarchy would be mass protection due to the existence of some approximately conserved chiral charges beyond the SM. An attractive possibility is that these chiral charges arise as a natural feature of the gauge symmetry group of the fundamental theory beyond the SM. This is the case in the family replicated gauge group model (also called the anti-grand unification model) [18,19]. The new chiral charges provide selection rules forbidding the transitions between the various left-handed and right-handed quark-lepton states, except for the top quark. In order to generate mass terms for the other fermion states, we have to introduce new Higgs fields, which break the symmetry group G of the fundamental theory down to the SM group. We also need suitable intermediate fermion states to mediate the forbidden transitions, which we take to be vector-like Dirac fermions

[2] See ref. [17] for a local chiral SU(3) family model with an alternative texture.

with a mass of order the fundamental scale M_F of the theory. In this way effective SM Yukawa coupling constants are generated [20], which are suppressed by the appropriate product of Higgs field VEVs measured in units of M_F. We assume that all the couplings in the fundamental theory are unsuppressed, i.e. they are all naturally of order unity.

The family replicated gauge group model is based on a non-simple non-supersymmetric extension of the SM with three copies of the SM gauge group—one for each family or generation. With the inclusion of three right-handed neutrinos, the gauge group becomes $G = (SMG \times U(1)_{B-L})^3$, where the three copies of the SM gauge group are supplemented by an abelian $(B - L)$ (= baryon number minus lepton number) gauge group for each family[3]. The gauge group G is the largest anomaly free group, transforming the known 45 Weyl fermions plus the three right-handed neutrinos into each other unitarily, which does *not* unify the irreducible representations under the SM gauge group. It is supposed to be effective at energies near to the Planck scale, $M_F = \Lambda_{Planck}$, where the i'th proto-family couples to just the i'th group factor $SMG_i \times U(1)_{B_i - L_i}$. The gauge group G is broken down by four Higgs fields W, T, ρ and ω, having VEVs about one order of magnitude lower than the Planck scale, to its diagonal subgroup:

$$(SMG \times U(1)_{B-L})^3 \to SMG \times U(1)_{B-L} \tag{29}$$

The diagonal $U(1)_{B-L}$ is broken down at the see-saw scale, by another Higgs field ϕ_{SS}, and the diagonal SMG is broken down to $SU(3) \times U(1)_{em}$ by the Weinberg-Salam Higgs field ϕ_{WS}.

Table 1. All $U(1)$ quantum charges of the Higgs fields in the $(SMG \times U(1)_{B-L})^3$ model.

	$y_1/2$	$y_2/2$	$y_3/2$	$(B-L)_1$	$(B-L)_2$	$(B-L)_3$
ω	$\frac{1}{6}$	$-\frac{1}{6}$	0	0	0	0
ρ	0	0	0	$-\frac{1}{3}$	$\frac{1}{3}$	0
W	0	$-\frac{1}{2}$	$\frac{1}{2}$	0	$-\frac{1}{3}$	$\frac{1}{3}$
T	0	$-\frac{1}{6}$	$\frac{1}{6}$	0	0	0
ϕ_{WS}	0	$\frac{2}{3}$	$-\frac{1}{6}$	0	$\frac{1}{3}$	$-\frac{1}{3}$
ϕ_{SS}	0	1	-1	0	2	0

The $(SMG \times U(1)_{B-L})^3$ gauge quantum numbers of the quarks and leptons are uniquely determined by the structure of the model and they include 6 chiral abelian charges—the weak hypercharge $y_i/2$ and $(B - L)_i$ quantum number for each of the three families, $i = 1, 2, 3$. With the choice of the abelian charges in Table 1 for the Higgs fields, it is possible to generate a good order of magnitude fit to the SM fermion masses, with VEVs of order $M_F/10$. In this fit, we do not attempt to guess the spectrum of superheavy fermions at the Planck scale, but

[3] The family replicated gauge groups $(SO(10))^3$ and $(E_6)^3$ have recently been considered by Ling and Ramond [21].

simply assume a sufficiently rich spectrum to mediate all of the symmetry breaking transitions in the various mass matrix elements. Then, using the quantum numbers of Table 1, the suppression factors are readily calculated as products of Higgs field VEVs measured in Planck units for all the fermion Dirac mass matrix elements[4], giving for example:

$$M_u \simeq \frac{\left\langle(\phi_{\rm WS})^\dagger\right\rangle}{\sqrt{2}}\begin{pmatrix} (\omega^\dagger)^3 W^\dagger T^2 & \omega\rho^\dagger W^\dagger T^2 & \omega\rho^\dagger (W^\dagger)^2 T \\ (\omega^\dagger)^4 \rho W^\dagger T^2 & W^\dagger T^2 & (W^\dagger)^2 T \\ (\omega^\dagger)^4 \rho & 1 & W^\dagger T^\dagger \end{pmatrix} \tag{30}$$

for the up quarks. Similarly the right-handed neutrino Majorana mass matrix is of order:

$$M_R \simeq \langle\phi_{\rm SS}\rangle\begin{pmatrix} (\rho^\dagger)^6 T^6 & (\rho^\dagger)^3 T^6 & (\rho^\dagger)^3 W^3 (T^\dagger)^3 \\ (\rho^\dagger)^3 T^6 & T^6 & W^3 (T^\dagger)^3 \\ (\rho^\dagger)^3 W^3 (T^\dagger)^3 & W^3 (T^\dagger)^3 & W^6 (T^\dagger)^{12} \end{pmatrix} \tag{31}$$

and the effective light neutrino mass matrix can be calculated from the Dirac neutrino mass matrix M_N and M_R using the see-saw formula [22]:

$$M_\nu = M_N M_R^{-1} M_N^T \tag{32}$$

In this way we obtain a good 5 parameter fit to the orders of magnitude of all the quark-lepton masses and mixing angles, as given in Table 2, actually even with the expected accuracy [23].

6 Conclusion

The hierarchical structure of the quark-lepton spectrum was emphasized and interpreted as due to the existence of a mass protection mechanism, controlled by approximately conserved chiral flavour quantum numbers beyond the SM. The family replicated gauge group model assigns a unique set of anomaly free gauge charges to the quarks and leptons. With an appropriate choice of quantum numbers for the Higgs fields, these chiral charges naturally generate a realistic set of quark-lepton masses and mixing angles. The top quark dominates the fermion mass matrices and we showed how the Multiple Point Principle can be used to predict the top quark and SM Higgs boson masses. We also discussed the lightest family mass generation model, which gives simple and compact formulae for all the CKM mixing angles in terms of the quark masses.

Acknowledgements

I should like to thank my collaborators Jon Chkareuli, Holger Bech Nielsen and Yasutaka Takanishi for many discussions.

[4] For clarity we distinguish between Higgs fields and their hermitian conjugates.

Table 2. Best fit to quark-lepton mass spectrum. All masses are running masses at 1 GeV except the top quark mass which is the pole mass.

	Fitted	Experimental
m_u	4.4 MeV	4 MeV
m_d	4.3 MeV	9 MeV
m_e	1.6 MeV	0.5 MeV
m_c	0.64 GeV	1.4 GeV
m_s	295 MeV	200 MeV
m_μ	111 MeV	105 MeV
M_t	202 GeV	180 GeV
m_b	5.7 GeV	6.3 GeV
m_τ	1.46 GeV	1.78 GeV
V_{us}	0.11	0.22
V_{cb}	0.026	0.041
V_{ub}	0.0027	0.0035
$\Delta m^2_\odot$	9.0×10^{-5} eV^2	5.0×10^{-5} eV^2
Δm^2_{atm}	1.7×10^{-3} eV^2	2.5×10^{-3} eV^2
$\tan^2\theta_\odot$	0.26	0.34
$\tan^2\theta_{atm}$	0.65	1.0
$\tan^2\theta_{chooz}$	2.9×10^{-2}	$\lesssim 2.6 \times 10^{-2}$

References

1. Particle Data Group, K. Hagiwara et al., Phys. Rev. D **66**, 010001 (2002).
2. M.C. Gonzales-Garcia and Y. Nir, Rev. Mod. Phys. **75**, 345 (2003) [arXiv:hep-ph/0202058].
3. C.D. Froggatt and H.B. Nielsen, These Proceedings.
4. N. Cabibbo, L. Maiani, G. Parisi and R. Petronzio, Nucl. Phys. B **158**, 295 (1979); M. Lindner, M. Sher and H.W. Zaglauer, Phys. Lett. B **228**, 139 (1989); M. Sher, Phys. Rep. **179**, 273 (1989).
5. M. Lindner, Z. Phys. C **31**, 295 (1986).
6. D.L. Bennett, C.D. Froggatt and H.B. Nielsen, Proceedings of the 27th International Conference on High Energy Physics, p. 557, ed. P. Bussey and I. Knowles (IOP Publishing Ltd, 1995); Perspectives in Particle Physics '94, p. 255, ed. D. Klabučar, I. Picek and D. Tadić (World Scientific, 1995) [arXiv:hep-ph/9504294].
7. S.W. Hawking, Phys. Rev. D **37**, 904 (1988); S. Coleman, Nucl. Phys. B **307**, 867 (1988); Nucl. Phys. B **310**, 643 (1988); T. Banks, Nucl. Phys. **309**, 493 (1988).
8. C.D. Froggatt and H.B. Nielsen, Phys. Lett. B **368**, 96 (1996) [arXiv:hep-ph/9511371].
9. J.A. Casas, J.R. Espinosa and M. Quiros Phys. Lett. B **342**, 171 (1995) [arXiv:hep-ph/9409458].
10. C.D. Froggatt, H.B. Nielsen and Y. Takanishi, Phys. Rev. D **64**, 113014 (2001) [arXiv:hep-ph/0104161].
11. C.D. Froggatt and H.B. Nielsen, arXiv:hep-ph/0308144.
12. H. Fritzsch, Phys. Lett. **B70**, 436 (1977).
13. P. Ramond, R.G. Roberts and G.G. Ross, Nucl. Phys. B **406**, 19 (1993) [arXiv:hep-ph/9303320].
14. J.L. Chkareuli and C.D. Froggatt, Phys. Lett. B **B50**, 158 (1999) [arXiv:hep-ph/9812499].

15. K. Matsuda, T. Fukuyama and H. Nishiura, Phys. Rev. D **60**, 013006 (1999) [arXiv:hep-ph/9902385].
16. J.L. Chkareuli, C.D. Froggatt and H.B. Nielsen, Nucl. Phys. B **626**, 307 (2002) [arXiv:hep-ph/0109156].
17. S.F. King, and G.G. Ross, Phys. Lett. B **574**, 239 (2003) [arXiv:hep-ph/0307190].
18. C.D. Froggatt, H.B. Nielsen and D.J. Smith, Phys. Lett. B **385**, 150 (1996) [arXiv:hep-ph/9607250]; C.D. Froggatt, M. Gibson, H.B. Nielsen and D.J. Smith, Int. J. Mod. Phys. A **13** 5037 (1998) [arXiv:hep-ph/9706212].
19. C.D. Froggatt, H.B. Nielsen and Y. Takanishi, Nucl. Phys. B **631**, 285 (2002) [arXiv:hep-ph/0201152]; H.B. Nielsen and Y. Takanishi, Phys. Lett. B **543**, 249 (2002) [arXiv:hep-ph/0205180].
20. C.D. Froggatt and H.B. Nielsen, Nucl. Phys. B **147**, 277 (1979).
21. F.S. Ling and P. Ramond, Phys. Rev. D **67**, 115010 (2003) [arXiv:hep-ph/0302264].
22. M. Gell-Mann, P. Ramond and R. Slansky in *Supergravity*, p. 315; Eds. P. van Nieuwenhuizen and D. Z. Freedman (North Holland, New York, 1980);
    T. Yanagida, Proc. of the Workshop on Unified Theory and Baryon Number in the Universe 1979, p. 95; Eds. A. Sawada and H. Sugawara (KEK, Tsukuba-Gu, Ibaraki-ken, Japan, 1979);
    R.N. Mohapatra and G. Senjanovic, Phys. Rev. Lett. **44** (1980).
23. C.D. Froggatt, H.B. Nielsen and D. Smith, Phys. Rev. D **65**, 095007 (2002) [arXiv:hep-ph/0108262].

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard model *(p. 184)*
Portorož, Slovenia, July 12-17, 2003

How to Approach Quantum Gravity – Background Independence in $1+1$ Dimensions

Daniel Grumiller*** and Wolfgang Kummer†

Institute f. Theor. Physics
Vienna University of Technology
Wiedner Hauptstr. 8-10, A-1040 Vienna, Austria

Abstract. The application of quantum theory to gravity is beset with many technical and conceptual problems. After a short tour d'horizon of recent attempts to master those problems by the introduction of new approaches, we show that the aim, a background independent quantum theory of gravity, can be reached in a particular area, 2d dilaton quantum gravity, *without* any assumptions beyond standard quantum field theory.

1 Introduction

It has been realized for some time that a merging of quantum theory with Einstein's theory of general relativity[1] (GR) is necessitated by consistency arguments. In *Gedankenexperimenten* the interaction of a classical gravitational wave with a quantum system inevitably leads to contradictions [3]. Arguments of this type are important because no relevant experimental data are available – we are very far from the quantum gravity analogue of the Balmer series.

On the other hand, when a quantum theory (QT) of gravity is developed along usual lines, one is confronted with a fundamental problem, from which many other (secondary) difficulties can be traced. The crucial difference to quantum field theory (QFT) in flat space is the fact that the variables of gravity exhibit a dual role, they are fields living on a manifold which is determined by themselves, "stage" and "actors" coincide. But there exist also numerous other problems: the time variable, an object with special properties already in QT, in GR appears on an equal footing with the space coordinates ("problem of time" which manifests itself in many disguises); the information paradox [4]; perturbative non-renormalizability [5] etc.

In section 2 we discuss some key-points regarding the definition of physical observables in QFT and the ensuing ones in quantum gravity (QGR). Then we critically mention some "old" and "new" approaches to QGR (section 3) from a strictly quantum field theorist's point of view. Finally we give some highlights

*** grumil@hep.itp.tuwien.ac.at
† wkummer@tph.tuwien.ac.at
[1] Several reviews on quantum gravity have emerged at the turn of the millennium, cf. e.g. [1,2].

on the "Vienna approach" to 2d dilaton quantum gravity with matter, including a new result (within that approach) on entropy corrections which is in agreement with the one found in literature (section 4). In that area which contains also models with physical relevance (e.g. spherically reduced gravity) the application of just the usual concepts of (even nonpertubative!) QFT lead to very interesting consequences [6] which allow physical interpretations in terms of "solid" traditional QFT observables.

2 Observables

2.1 Cartan variables in GR

Physical observables in the sense used here are certain functionals of the field variables which are directly accessible to experimental measurements.

The metric g in GR can be considered as a "derived" field variable

$$g = e^a \otimes e^b \eta_{ab}, \tag{1}$$

because it is the direct product of the dual basis one forms[2] $e^a = e^a_\mu \, dx^\mu$ contracted with the flat local Lorentz metric η_{ab} which is used to raise and lower "flat indices" denoted by Latin letters ($\eta = \mathrm{diag}(1, -1, -1, -1, \ldots .)$, $x^\mu = \{x^0, x^i\}$). Local Lorentz invariance leads to the "covariant derivative" $D^a{}_b = \delta^a_b d + \omega^a{}_b$ with a spin connection 1-form $\omega^a{}_b$ as a gauge field. Its antisymmetry $\omega^{ab} = -\omega^{ba}$ implies metricity. Thanks to the Bianchi identities all covariant tensors relevant for constructing actions in even dimensions can be expressed in terms of e^a, the curvature 2-form $R^{ab} = (D\omega)^{ab}$ and the torsion 2-form $T^a = (De)^a$. For nonvanishing torsion the affine connection $\Gamma_{\mu\nu}{}^\rho = E^\rho_a \, (D_\mu e)^a_\nu$, expressed in terms of components e^a_μ and of its inverse E^ρ_a, besides the usual Christoffel symbols also contains a contorsion term in $\Gamma_{(\mu\nu)}{}^\rho$, whereas $\Gamma_{[\mu\nu]}{}^\rho$ are the components of torsion. Einstein gravity in d=4 dimensions postulates vanishing torsion $T^a = 0$ so that $\omega = \omega(e)$. This theory can be derived from the Hilbert action (G_N is Newton's constant; dS space results for nonvanishing cosmological constant Λ from the replacement $R^{ab} \to R^{ab} - \frac{4}{3}\Lambda e^a \wedge e^b$)

$$L_{(H)} = \frac{1}{16\pi G_n} \int_{\mathcal{M}_4} R^{ab} \wedge e^c \wedge e^d \epsilon_{abcd} + L_{(\mathrm{matter})}. \tag{2}$$

Because of the "Palatini mystery", independent variation of $\delta\omega$ yields $T^a = 0$, whereas δe produces the Einstein equations.

Instead of working with the metric (1) the "new" approaches [8] are based upon a gauge field related to ω^{ab}

$$A^{ab} = \frac{1}{2}\left(\omega^{ab} - \frac{\gamma}{2}\epsilon^{ab}{}_{cd}\,\omega^{cd}\right). \tag{3}$$

[2] For details on gravity in the Cartan formulation we refer to the mathematical literature, e.g. [7]

The Barbero-Immirzi parameter γ [9] is an arbitrary constant. The extension to complex gravity ($\gamma = i$) makes A^a a self-adjoint field and transforms the Einstein theory into the one of an $SU(2)$ gauge field

$$A_i^{\underline{a}} = \epsilon^{0\,\underline{a}}{}_{\underline{b}\,\underline{c}}\, A_i^{\underline{b}\,\underline{c}}, \tag{4}$$

where the index $\underline{a} = 1, 2, 3$. This formulation is the basis of loop quantum gravity and spin foam models (see below).

2.2 Observables in classical GR

The exploration of the global properties of a certain solution of (2), its singularity structure etc., is only possible by means of the introduction of an additional test field, most simply a test particle with action

$$\begin{aligned} L_{(\text{test})} &= -m_0 \int |ds|\,, \\ ds^2 &= g_{\mu\nu}\left(x\left(\tau\right)\right) \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau}, \end{aligned} \tag{5}$$

which is another way to incorporate Einstein's old proposal [10] of a "net of geodesics". The path $x^{\mu}(\tau)$ is parameterized by the affine parameter τ (actually only timelike or lightlike $ds^2 \geq 0$ describes the paths of a physical particle).

It is not appreciated always that the global properties of a manifold are *defined* in terms of a specific device like (5). Whereas the usual geodesics derived from (5) depend on $g_{\mu\nu}$ through the Christoffel symbols only, e.g. in the case of torsion also the contorsion may contribute ("autoparallels") in the affine connection; spinning particles "feel" the gravimagnetic effect etc. As a consequence, when a field dependent transformation of the gravity variables is performed (e.g. a conformal transformation from a "Jordan frame" to an "Einstein frame" in Jordan-Brans-Dicke [11] theory) the action of the device must be transformed in the same way.

2.3 Observables in QFT

In flat QFT one starts from a Schrödinger equation, dependent on field operators and, proceeding through Hamiltonian quantization to the path integral, the experimentally accessible observables are the elements of the S-matrix, or quantities expressible by those.[3] It should be recalled that the properly defined renormalized S-matrix element obtains by amputation of external propagators in the related Green function, multiplication with polarizations and with a square root of the wave function renormalization constant, taking the mass-shell limit.

In gauge theories one encounters the additional problem of gauge-dependence, i.e. the dependence on some gauge parameter β introduced by generic

[3] Note that ordinary quantum mechanics and its Schrödinger equation appear as the nonrelativistic, weak coupling limit of the Bethe-Salpeter equation of QFT [12]. Useful notions like eigenvalues of Hermitian operators, collapse of wave functions etc. are not basic concepts in this more general frame (cf. footnote 2 in ref. [13]).

gauge fixing. Clearly the S-matrix elements must be and indeed are [13] independent of β. But other objects, in particular matrix-elements of gauge invariant operators $\mathcal{O}_A$, depend on β. In addition, under renormalization they mix with operators $\tilde{\mathcal{O}}_{\tilde{A}}$ of the same "twist" (dimension minus spin) which depend on Faddeev-Popov ghosts [14] and are not gauge-invariant:

$$\begin{aligned} \mathcal{O}_A^{(ren)} &= Z_{AB}\mathcal{O}_B + Z_{A\tilde{B}}\tilde{\mathcal{O}}_{\tilde{B}} \\ \tilde{\mathcal{O}}_{\tilde{A}}^{(ren)} &= \qquad\qquad Z_{\tilde{A}\tilde{B}}\tilde{\mathcal{O}}_{\tilde{B}} \end{aligned} \tag{6}$$

The contribution of such operators to the S-matrix element (sic!) of e.g. the scaling limit for deep inelastic scattering [15] of leptons on protons [16] occurs only through the anomalous dimensions ($\propto \partial Z_{AB}/\partial\Lambda$ for a regularisation cut-off Λ). And those objects, also thanks to the triangular form of (6), are gauge-independent!

In flat QFT, as well as in QGR, the (gauge invariant) "Wilson loop"

$$W_{(\mathcal{C})} = \mathrm{Tr}\, P \exp\left(i \oint_{\mathcal{C}} A_{\mu\, dx^\mu}\right), \tag{7}$$

parameterized by a path ordered closed curve $\mathcal{C}$, often is assumed to play an important role. In covariant gauges it is multiplicatively renormalizable with the renormalization constant depending on the length of $\mathcal{C}$, the UV cut-off and eventual cusp-angles in $\mathcal{C}$ [17]. Still the relation to experimentally observable quantities (should one simply drop the renormalization constant or proceed [13] as for an S-matrix?) is unclear. Worse, for lightlike axial gauges $(nA) = 0$ $\left(n^2 = 0\right)$ multiplicative renormalization is not applicable [18]. Then, only for a matrix element of (7) between "on-shell gluons", this type of renormalization is restored. Still the renormalization constant is different from covariant gauge, except for the anomalous dimension derived from it (cf. precisely that feature of operators in deep inelastic scattering).

3 Approaches to QGR

"Old" QGR worked with a separation of the two aspects of gravity variables by the decomposition of the metric

$$g_{\mu\nu} = g_{\mu\nu}^{(0)} + h_{\mu\nu}, \tag{8}$$

which consists of a (fixed) classical background $g_{\mu\nu}^{(0)}$ ("stage") with small quantum fluctuations $h_{\mu\nu}$ ("actors"). The "observable" (to be tested by a classical device) would be the effective matrix $g_{\mu\nu}^{(\mathrm{eff})} = g_{\mu\nu}^{(0)} + < h_{\mu\nu} >$. Starting computations from the action (2) one finds that an ever increasing number of counterterms is necessary. They are different from the terms in the Lagrangian $\mathcal{L} = \sqrt{-g}\, R/(16\pi G_N)$ in (2). This is the reason why QGR is called (perturbatively) "nonrenormalizable" [5]. Still, at energies $E \ll (G_N)^{-1/2}$, i.e. much below the Planck mass scale $m_{Pl} \propto (G_N)^{-1/2}$, such calculations can be meaningful in the sense of an "effective low energy field theory" [19], irrespective of the fact that

(perhaps by embedding gravity into string theory) by inclusion of further fields at higher energy scales (Planck scale), QGR may become renormalizable. Of course, such an approach even when it is modified by iterative inclusion of $< h_{\mu\nu} >$ into $g^{(0)}_{\mu\nu}$ etc. – which is technically quite hopeless – completely misses inherent background independent effects, i.e. effects when $g^{(0)}_{\mu\nu} = 0$.

One could think also of applying nonperturbative methods developed in numerical lattice calculations for QCD. However, there are problems to define the Euclidean path integral for that, because the Euclidean action is not bounded from below (as it is the case in QCD) [20].

The quantization of gravity which – at least in principle – avoids background dependence is based upon the ADM approach to the Dirac quantization of the Hamiltonian [21]. Space-time is foliated by a sequence of three dimensional space-like manifolds $\sum_3$ upon which the variables $g_{ij} = q_{ij}$ and associated canonical momenta π_{ij} live. The constraints associated to the further variables lapse (N_0) and shift (N_i) in the Hamiltonian density

$$\mathcal{H} = N_0 \, H^0 \, (q, \pi) + N_i \, H^i \, (q, \pi) \tag{9}$$

are primary ones. The Poisson brackets of the secondary constraints H^μ closes. H^i generates diffeomorphisms on $\sum_3$. In the quantum version of (9) the solutions of the Wheeler-deWitt equation involving the Hamiltonian constraint

$$\int\limits_{\sum_3} H^0 \left(q, \frac{\delta}{i\delta q} \right) \mid \psi >= 0 \tag{10}$$

formally would correspond to a nonperturbative QGR. Apart from the fact that it is extremely difficult to find a general solution to (10) there are several basic problems with a quantum theory based upon that equation (e.g. no Hilbert space $\mid \psi >$ can be constructed, no preferred time foliation exists with ensuing inequivalent quantum evolutions [22], problems with usual "quantum causality" exist, the "axiom" that fields should commute at space like distances does not hold etc.). A restriction to a finite number of degrees of freedom ("mini superspace") [23] or infinite number of degrees of freedom (but still less than the original theory – so-called "midi superspace") [24] has been found to miss essential features.

As all physical states $\mid \psi >$ must be annihilated by the constraints H^μ, a naive Schrödinger equation involving the Hamiltonian constraint H^0,

$$i\hbar \frac{\partial \mid \psi >}{\partial t} = H^0 \mid \psi >= 0, \tag{11}$$

cannot contain a time variable ("problem of time"). A kind of Schrödinger equation can be produced by the definition of a "time-function" $T\,(q, \pi, x)$, at the price of an even more complicated formalism [25] with quite ambiguous results – and the problem, how to connect those with "genuine" observables. All these problems are aggravated, when one tries to first eliminate constraints by solving them explicitly before quantization. In this way, clearly part of the quantum fluctuations are eliminated from the start. As a consequence different quantum theories, constructed in this way, are not equivalent.

The "new" gravities (loop quantum gravity, spin foam models) reformulate the quantum theory of space-time by the introduction of novel variables, based upon the concept of Wilson loops (7) applied to the gauge-field (4). The operator

$$U(s_1, s_2) = \mathrm{Tr}\, P\, \exp\left(i \int_{s_1}^{s_2} ds \frac{dx^i}{ds} A_i \right) \tag{12}$$

defines a holonomy. It is generalized by inserting further invariant operators at intermediate points between s_1 and s_2 . From such holonomies a spin network can be created which represents spacetime (in the path integral it is dubbed "spin foam").

These approaches claim several successes [2]. Introducing as a basis diffeomorphism equivalence classes of "labeled graphs" a finite Hilbert space can be constructed and some solutions of the Wheeler-deWitt equation (10) have been obtained. The methods introduce a "natural" coarse graining of space-time which implies a UV cutoff. "Small" gravity around certain states leads in those cases to corresponding linearized Einstein gravity.

However, despite of very active research in this field a number of very serious open questions persists: The Hamiltonian constructed from spin networks does not lead to massless excitations (gravitons) in the classical limit. The Barbero-Immirzi parameter γ has to be fixed by the requirement of a "correct" Bekenstein-Hawking entropy for the Black Hole. The most severe problem, however, is the one of observables. By some researchers in this field it has been claimed that by "proper gauge fixing" (!) area and volume can be obtained as quantized "observables", which is a contradiction in itself from the point of view of QFT. We must emphasize too that also in an inherently UV regularized theory (finite) renormalization remains an issue to be dealt with properly. Also the fate of S-matrix elements, which play such a central role as the proper observables in QFT, is completely unclear in these setups.

Embedding QGR into (super-)string theory [26] does not remove the key problems related to the dual role of the metric. Gravity may well be a string excitation in a string/brane world of 10-11 dimensions, possibly a finite theory of everything. Nevertheless, at low energies Einstein gravity (eventually plus an antisymmetric B-field) remains the theory for which computations must be performed.[4] Unfortunately, the proper choice (let alone the derivation) of a string vacuum in our d=4 space-time is an unsolved problem.

Many other approaches exist, including noncommutative geometry, twistors, causal sets, 3d approaches, dynamical triangulations, Regge calculus etc., each of which has certain attractive features and difficulties (cf. e.g. [2] and refs. therein).

To us all these "new" approaches appear as – very ingenious – attempts to bypass the technical problems of directly applying standard QFT to gravity – without a comprehensive solution of the main problems of QGR being in sight.

[4] It should be noted that the now widely confirmed astronomical observations of a positive cosmological constant [27] (if it is a constant and not a "quintessence" field in a theory of type [11]) precludes immediate application of supersymmetry (supergravity) in string theory, because only AdS space is compatible with supergravity [28].

Thus the main points of a "minimal" QFT for gravity should be based upon "proven concepts" of QFT with a point of departure characterizing QGR as follows:

(a) QGR is an "effective" low energy theory and therefore need not be renormalizable to all orders.
(b) QGR is based upon classical Einstein (-dS) gravity with usual variables (metric or Cartan variables).
(c) At least the quantization of geometry must be performed in a background independent (nonperturbative) way.
(d) Absolutely "safe" quantum observables are only the S-matrix elements of QFT $< f \mid S \mid i >$, where initial state $\mid i >$ and final state $< f \mid$ are defined only when those states are realized as Fock states of particles in a (at least approximate) flat space environment. In certain cases it is permissible to employ a semi-classical approach: expectation values of quantum corrections may be added to classical geometric variables, and a classical computation is then based on the effective variables, obtained in this way.

Clearly item (d) by construction excludes any application to quantum cosmology, where $\mid i >$ would be the (probably nonexistent) infinite past before the Big Bang.

Obviously the most difficult issue is (c). We describe in the following section how gravity models in d=2 (e.g. spherically reduced gravity) permit a solution of just that crucial point, leading to novel results.

4 "Minimal" QGR in 1+1 dimensions

4.1 Classical theory: first order formulation

In the 1990s the interest in dilaton gravity in d=2 was rekindled by results from string theory [29], but it existed as a field on its own more or less since the 1980s [30]. For a review on dilaton gravity ref. [6] may be consulted. For sake of self-containment the study of dilaton gravity will be motivated briefly from a purely geometrical point of view.

The notation of ref. [6] is used: e^a is the zweibein one-form, $\epsilon = e^+ \wedge e^-$ is the volume two-form. The one-form ω represents the spin-connection $\omega^a{}_b = \varepsilon^a{}_b \omega$ with the totally antisymmetric Levi-Civitá symbol ε_{ab} ($\varepsilon_{01} = +1$). With the flat metric η_{ab} in light-cone coordinates ($\eta_{+-} = 1 = \eta_{-+}, \eta_{++} = 0 = \eta_{--}$) the torsion 2-form reads $T^\pm = (d \pm \omega) \wedge e^\pm$. The curvature 2-form $R^a{}_b$ can be presented by the 2-form R defined by $R^a{}_b = \varepsilon^a{}_b R$, $R = d \wedge \omega$. Signs and factors of the Hodge-$*$ operation are defined by $*\epsilon = 1$.

Since the Hilbert action $\int_{\mathcal{M}_2} R \propto (1-g)$ yields just the Euler number for a surface with genus g one has to generalize it appropriately. The simplest idea is to introduce a Lagrange multiplier for curvature, X, also known as "dilaton field", and an arbitrary potential thereof, $V(X)$, in the action $\int_{\mathcal{M}_2} (XR + \epsilon V(X))$. In particular, for $V \propto X$ the Jackiw-Teitelboim model emerges [30]. Having introduced curvature it is natural to consider torsion as well. By analogy the first

order gravity action [31]

$$L^{(1)} = \int_{\mathcal{M}_2} (X_a T^a + XR + \epsilon \mathcal{V}(X^a X_a, X)) \tag{13}$$

can be motivated where X_a are the Lagrange multipliers for torsion. It encompasses essentially all known dilaton theories in 2d, also known as Generalized Dilaton Theories (GDT). Spherically reduced gravity (SRG) from d=4 corresponds to $\mathcal{V} = -X^+X^-/(2X) - \text{const}$.

Without matter there are no physical propagating degrees of freedom, which is advantageous mathematically but not very attractive from a physical point of view. Thus, in order to describe scattering processes matter has to be added. The simplest way is to consider a massless Klein-Gordon field ϕ,

$$L^{(m)} = \frac{1}{2} \int_{\mathcal{M}_2} F(X)\, d\phi \wedge *d\phi \,, \tag{14}$$

with a coupling function $F(X)$ depending on the dilaton (for dimensionally reduced theories typically $F \propto X$ holds).

4.2 Quantum theory: Virtual Black Holes

It turned out that even in the presence of matter an exact path integration of all geometric quantities is possible for all GDTs, proceeding along well established paths of QFT[5] [32].

The effective theory obtained in this way solely depends on the matter fields in which it is nonlocal and non-polynomial. Already at the level of the (nonlocal) vertices of matter fields, to be used in a systematic perturbative expansion in terms of Newton's constant, a highly nontrivial and physically intriguing phenomenon can be observed, namely the so-called "virtual black hole" (VBH). This notion originally has been introduced by S. Hawking [33], but in our recent approach the VBH for SRG emerges naturally in Minkowski signature space-time, without the necessity of additional *ad hoc* assumptions.

For non-minimally coupled scalars the lowest order S-matrix indeed exhibited interesting features: forward scattering poles, monomial scaling with energy, CPT invariance, and pseudo-self-similarity in its kinematic sector [34].

It was possible to reconstruct geometry self-consistently from a (perturbative or, if available, exact) solution of the effective theory. For the simplest case of four-point tree-graph scattering the corresponding Carter-Penrose (CP) diagram is presented in Fig. 1. It is non-local in the sense that it depends not only on one set of coordinates but on two. This was a consequence of integrating out geometry

[5] We mention just a few technical details: no ordering ambiguities arise, the (nilpotent) BRST charge is essentially the same as for Yang-Mills theory (despite of the appearance of nonlinearities in the algebra of the first class secondary constraints), the gauge fixing fermion is chosen such that "temporal" gauge is obtained, the Faddeev-Popov determinant cancels after integrating out the "unphysical" sector, and "natural" boundary conditions cannot be imposed on the fields, so one has to be careful with the proper treatment of the boundary.

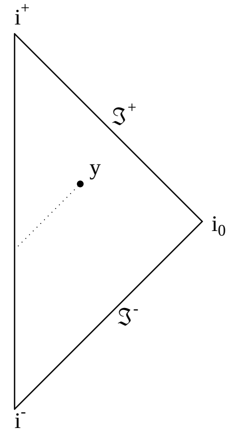

Fig. 1. CP diagram of the VBH

non-perturbatively. For each choice of y (one of the two sets of coordinates) it is possible to draw an ordinary CP-diagram. The non-trivial part of our effective geometry (i.e. the VBH) is concentrated on the light-like cut. For SRG the ensuing line-element has Sachs-Bondi form

$$(ds)^2 = 2drdu + \left(1 - \frac{2m(u,r)}{r} - a(u,r)r + d(u,r)\right)(du)^2 \,, \tag{15}$$

with m, a and d being localized[6] on the cut $u = u_0$ with compact support $r < r_0$. These quantities depend on the second set of coordinates u_0, r_0.

One should not take the effective geometry at face value – this would be like over-interpreting the role of virtual particles in a loop diagram. It is a nonlocal entity and one still has to "sum" (read: integrate) over all possible geometries of this type in order to obtain the nonlocal vertices and the scattering amplitude. Nonetheless, the simplicity of this geometry and the fact that all possible configurations are summed over are nice features of this picture. Moreover, all VBH geometries coincide asymptotically and differ only very little from each other in the asymptotic region. This observation allows for the following interpretation: the boundaries of the diagram, $\mathcal{I}^{\pm}$ and i^0, behave in a classical way[7] (thus enabling one to construct an ordinary Fock space like in fixed background QFT), but the more one zooms into the geometry the less classical it becomes. The situation seems to be quite contrary to Kuchař's proposal of geometrodynamics[8] of BHs: while we have fixed boundary conditions for the target space coordinates

[6] The localization of "mass" and "Rindler acceleration" on a light-like cut is not an artifact of an accidental gauge choice, but has a physical interpretation in terms of the Ricci-scalar. Certain parallels to Hawking's Euclidean VBHs can be observed, but also essential differences. The main one is our Minkowski signature which we deem to be a positive feature.

[7] Clearly the imposed boundary conditions play a crucial role in this context. They produce effectively a fixed background, but only at the boundary.

[8] This approach considers only the matterless case and thus a full comparison to our results is not possible.

(and hence a fixed ADM mass) but a "smeared geometry" (in the sense that a continuous spectrum of asymptotically equivalent VBHs contributes to the S-matrix), Kuchař encountered a "smeared mass" (obeying a Schrödinger equation) but an otherwise fixed geometry [35].

Qualitatively it is clear what has to be done in order to obtain the S-matrix[9]: Take all possible VBHs of Fig. 1 and sum them coherently with proper weight factors and suitably attached external legs of scalar fields. This had been done quantitatively in a straightforward but rather lengthy calculation for gravitational scattering of s-waves in the framework of SRG, the result of which yielded the lowest order tree-graph S-matrix for ingoing modes with momenta q, q' and outgoing ones k, k',

$$T(q,q';k,k') = -\frac{i\kappa\delta\,(k+k'-q-q')}{2(4\pi)^4|kk'qq'|^{3/2}}E^3\tilde{T}\,, \tag{16}$$

with the total energy $E = q + q'$, $\kappa = 8\pi G_N$,

$$\tilde{T}(q,q';k,k') := \frac{1}{E^3}\left[\Pi \ln\frac{\Pi^2}{E^6} + \frac{1}{\Pi}\sum_{p\in\{k,k',q,q'\}} p^2 \ln\frac{p^2}{E^2}\left(3kk'qq' - \frac{1}{2}\sum_{r\neq p}\sum_{s\neq r,p}\left(r^2s^2\right)\right)\right], \tag{17}$$

and the momentum transfer function $\Pi = (k+k')(k-q)(k'-q)$. The interesting part of the scattering amplitude is encoded in the scale independent factor $\tilde{T}$. The forward scattering poles occurring for $\Pi = 0$ should be noted.

It is possible to generalize the VBH phenomenon to arbitrary GDTs with matter as well as most of its properties (for instance, the CP-diagram, CPT invariance and the role played in the S-matrix) [37].

4.3 New results and outlook

Recently quantum corrections to the specific heat of the dilaton BH have been calculated by applying the quantization method discussed above [38]. The result is $C_s := dM/dT = 96\pi^2 M^2/\lambda^2$, where λ is the scale parameter of the theory. Thus, in that particular case quantum corrections lead to a stabilization of the system. The mass of the BH is found to be decreasing according to

$$M(u) \approx M_0 - \frac{\pi}{6}(T_H^0)^2(u-u_0) - \frac{\lambda}{24\pi}\ln\frac{M(u)}{M_0} + \mathcal{O}\left(\frac{\lambda}{M(u)}\right)\,. \tag{18}$$

The first term is the ADM mass, the second term corresponds to a linear decrease due to the (in leading order) constant Hawking flux and the third term provides the first nontrivial correction.

[9] The idea that BHs must be considered in the S-matrix together with elementary matter fields has been put forward some time ago [36]. The approach [34] reviewed here, for the first time allowed to derive (rather than to conjecture) the appearance of the BH states in the quantum scattering matrix of gravity.

Applying simple thermodynamical methods[10] ($dS = C_s\, dT/T$) and exploiting the quantum corrected mass/temperature relation $T/T_0 = 1 - \lambda/(48\pi M)$ it is possible to calculate also entropy corrections:

$$S = S_0 - \frac{1}{24}\ln S_0 + \mathcal{O}(1)\,, \quad S_0 := \frac{2\pi M}{\lambda} = 2\pi X|_{\text{horizon}} \tag{19}$$

The logarithmic behavior is in qualitative agreement with the one found in the literature by various methods [40]; the prefactor 1/24 coincides with [41].

An extension of the results obtained in the first order formulation to dilaton supergravity is straightforward in principle but somewhat tedious in detail. It permitted, among other results, to obtain for the first time a full solution of dilaton supergravity [42].

All these exciting applications indicate that the strict application of standard QFT concepts to gravity (at least in d=2 or in models dimensionally reduced to d=2) shows great promise.

Acknowledgement

This work has been supported by project P-14650-TPH of the Austrian Science Foundation (FWF) and by EURESCO. We are grateful to the organizers of the workshop “What comes beyond the Standard Model?” for an enjoyable meeting in Slovenia. We thank M. Bojowald and O. Zaslavskii for useful discussions at the Erwin Schrödinger Institut in Vienna, L. Bergamin for collaboration on 2d dilaton supergravity and D. Vassilevich for a long time collaboration on 2d dilaton gravity.

References

1. G. T. Horowitz, (1996), gr-qc/9604051; C. Rovelli, Living Rev. Rel. **1**, 1 (1998), gr-qc/9710008; Notes for a brief history of quantum gravity, in *9th Marcel Grossmann Meeting On Recent Developments In Theoretical And Experimental General Relativity, Gravitation And Relativistic Field Theories*, 2000, arXiv:gr-qc/0006061.
2. S. Carlip, Rept. Prog. Phys. **64**, 885 (2001), arXiv:gr-qc/0108040; L. Smolin, (2003), hep-th/0303185.
3. K. Eppley and E. Hannah, Found. Phys. **7**, 51 (1977).
4. T. Banks, Nucl. Phys. Proc. Suppl. **41**, 21 (1995), hep-th/9412131.
5. G. ’t Hooft and M. J. G. Veltman, Annales Poincare Phys. Theor. **A20**, 69 (1974).
6. D. Grumiller, W. Kummer, and D. V. Vassilevich, Phys. Rept. **369**, 327 (2002), hep-th/0204253.
7. M. Nakahara, *Geometry, Topology and Physics* (IOP Publishing, Bristol, 1990).
8. A. Sen, Phys. Lett. **B119**, 89 (1982); A. Ashtekar, Phys. Rev. Lett. **57**, 2244 (1986); Phys. Rev. **D36**, 1587 (1987); C. Rovelli, Class. Quant. Grav. **8**, 1613 (1991).
9. J. F. Barbero, Phys. Rev. **D51**, 5507 (1995), gr-qc/9410014; G. Immirzi, Class. Quant. Grav. **14**, L177 (1997), gr-qc/9612030; Nucl. Phys. Proc. Suppl. **57**, 65 (1997), gr-qc/9701052.

[10] The review [39] may be consulted in this context.

10. A. Einstein, Annalen Phys. **49**, 769 (1916).
11. M. Fierz, Helv. Phys. Acta **29**, 128 (1956); P. Jordan, Z. Phys. **157**, 112 (1959); C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
12. E. E. Salpeter and H. A. Bethe, Phys. Rev. **84**, 1232 (1951); N. Nakanishi, Prog. Theor. Phys. Suppl. **43**, 1 (1969); R. Barbieri and E. Remiddi, Nucl. Phys. **B141**, 413 (1978); W. E. Caswell and G. P. Lepage, Phys. Rev. **A18**, 810 (1978).
13. W. Kummer, Eur. Phys. J. **C21**, 175 (2001), hep-th/0104123.
14. J. A. Dixon and J. C. Taylor, Nucl. Phys. **B78**, 552 (1974); J. A. Dixon, Nucl. Phys. **B99**, 420 (1975); H. Kluberg-Stern and J. B. Zuber, Phys. Rev. **D12**, 467 (1975); W. Konetschny and W. Kummer, Nucl. Phys. **B124**, 145 (1977); S. D. Joglekar and B. W. Lee, Ann. Phys. **97**, 160 (1976).
15. J. I. Friedman, H. W. Kendall, and R. E. Taylor, Rev. Mod. Phys. **63**, 573 (1991).
16. D. J. Gross and F. Wilczek, Phys. Rev. **D8**, 3633 (1973); Phys. Rev. **D9**, 980 (1974).
17. A. M. Polyakov, Nucl. Phys. **B164**, 171 (1980); V. S. Dotsenko and S. N. Vergeles, Nucl. Phys. **B169**, 527 (1980); R. A. Brandt, F. Neri, and M.-a. Sato, Phys. Rev. **D24**, 879 (1981); J. Frenkel and J. C. Taylor, Nucl. Phys. **B246**, 231 (1984).
18. A. Andrasi, Eur. Phys. J. **C18**, 601 (2001), hep-th/9912138.
19. J. F. Donoghue, Phys. Rev. **D50**, 3874 (1994), gr-qc/9405057.
20. G. W. Gibbons and S. W. Hawking, editors, *Euclidean quantum gravity* (Singapore: World Scientific, 1993).
21. R. Arnowitt, S. Deser, and C. W. Misner, in *Gravitation: An Introduction to Current Research*, edited by L. Witten, Wiley, New York, 1962.
22. C. G. Torre and M. Varadarajan, Class. Quant. Grav. **16**, 2651 (1999), hep-th/9811222.
23. B. S. DeWitt, Phys. Rev. **160**, 1113 (1967); C. W. Misner, Phys. Rev. **186**, 1319 (1969).
24. K. Kuchař, Phys. Rev. **D4**, 955 (1971).
25. P. Bergmann and A. Komar, in *Recent developments in general relativity*, Pergamon, New York, 1962; C. Teitelboim, Phys. Lett. **B56**, 376 (1975); K. Kuchař, in *Quantum gravity 2: an Oxford symposium*, edited by C. Isham, R. Penrose, and D. Sciama, Clarendon Press, Oxford, 1975.
26. J. Polchinski, *String theory* (Cambridge University Press, 1998), Vol. 1: An Introduction to the Bosonic String; Vol. 2: Superstring Theory and Beyond.
27. A. G. Riess *et al.*, Astrophys. J. **560**, 49 (2001), astro-ph/0104455; A. G. Riess, The Publications of the Astronomical Society of the Pacific **112**, 1284 (2000); Supernova Search Team, A. G. Riess *et al.*, Astron. J. **116**, 1009 (1998), astro-ph/9805201; Supernova Cosmology Project, S. Perlmutter *et al.*, Astrophys. J. **517**, 565 (1999), arXiv:astro-ph/9812133.
28. P. Van Nieuwenhuizen, Phys. Rept. **68**, 189 (1981).
29. G. Mandal, A. M. Sengupta, and S. R. Wadia, Mod. Phys. Lett. **A6**, 1685 (1991); S. Elitzur, A. Forge, and E. Rabinovici, Nucl. Phys. **B359**, 581 (1991); E. Witten, Phys. Rev. **D44**, 314 (1991); R. Dijkgraaf, H. Verlinde, and E. Verlinde, Nucl. Phys. **B371**, 269 (1992); C. G. Callan, Jr., S. B. Giddings, J. A. Harvey, and A. Strominger, Phys. Rev. **D45**, 1005 (1992), hep-th/9111056.
30. B. M. Barbashov, V. V. Nesterenko, and A. M. Chervyakov, Teor. Mat. Fiz. **40**, 15 (1979); J. Phys. **A13**, 301 (1979); Theor. Math. Phys. **40**, 572 (1979); E. D'Hoker and R. Jackiw, Phys. Rev. **D26**, 3517 (1982); C. Teitelboim, Phys. Lett. **B126**, 41 (1983); E. D'Hoker, D. Freedman, and R. Jackiw, Phys. Rev. **D28**, 2583 (1983); E. D'Hoker and R. Jackiw, Phys. Rev. Lett. **50**, 1719 (1983); R. Jackiw, (1995), hep-th/9501016.
31. N. Ikeda and K. I. Izawa, Prog. Theor. Phys. **90**, 237–246 (1993), hep-th/9304012; N. Ikeda, Ann. Phys. **235** 435–464 (1994), hep-th/9312059; P. Schaller and T. Strobl, Mod. Phys. Lett. **A9** 3129–3136 (1994), hep-th/9405110.

32. W. Kummer and D. J. Schwarz, Nucl. Phys. **B382**, 171 (1992); F. Haider and W. Kummer, Int. J. Mod. Phys. **A9**, 207 (1994); W. Kummer, H. Liebl, and D. V. Vassilevich, Nucl. Phys. **B493**, 491 (1997), gr-qc/9612012; Nucl. Phys. **B544**, 403 (1999), hep-th/9809168; D. Grumiller, *Quantum dilaton gravity in two dimensions with matter*, PhD thesis, Technische Universität Wien, 2001, gr-qc/0105078.
33. S. W. Hawking, Phys. Rev. **D53**, 3099 (1996), hep-th/9510029.
34. D. Grumiller, W. Kummer, and D. V. Vassilevich, Nucl. Phys. **B580**, 438 (2000), gr-qc/0001038; P. Fischer, D. Grumiller, W. Kummer, and D. V. Vassilevich, Phys. Lett. **B521**, 357 (2001), gr-qc/0105034, Erratum ibid. **B532** (2002) 373; D. Grumiller, Int. J. Mod. Phys. **A 17**, 989-992 (2001), hep-th/0111138; Class. Quant. Grav. **19**, 997 (2002), gr-qc/0111097.
35. K. V. Kuchař, Phys. Rev. **D50**, 3961 (1994), arXiv:gr-qc/9403003.
36. G. 't Hooft, Int. J. Mod. Phys. **A11**, 4623 (1996), gr-qc/9607022.
37. D. Grumiller, W. Kummer, and D. V. Vassilevich, European Phys. J. **C30**, 135 (2003), hep-th/0208052.
38. D. Grumiller, W. Kummer, and D. V. Vassilevich, JHEP **07**, 009 (2003), hep-th/0305036.
39. R. M. Wald, Living Rev. Rel. **4**, 6 (2001), gr-qc/9912119.
40. R. B. Mann and S. N. Solodukhin, Nucl. Phys. **B523**, 293 (1998), hep-th/9709064; R. K. Kaul and P. Majumdar, Phys. Rev. Lett. **84**, 5255 (2000), gr-qc/0002040; S. Carlip, Class. Quant. Grav. **17**, 4175 (2000), gr-qc/0005017.
41. T. M. Fiola, J. Preskill, A. Strominger and S. P. Trivedi, Phys. Rev. D **50**, 3987 (1994), hep-th/9403137; R. C. Myers, Phys. Rev. D **50**, 6412 (1994), hep-th/9405162; J. D. Hayward, Phys. Rev. D **52**, 2239 (1995), gr-qc/9412065; O. B. Zaslavskii, Phys. Lett. **B375**, 43-46 (1996).
42. L. Bergamin and W. Kummer, (2003), hep-th/0306217; L. Bergamin, D. Grumiller, and W. Kummer, (2003), hep-th/0310006.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard model *(p. 197)*
Portorož, Slovenia, July 12-17, 2003

Hidden Spacetime Symmetries and Generalized Holonomy in M-theory⋆, ⋆⋆

Michael J. Duff⋆⋆⋆ and James T. Liu†

Michigan Center for Theoretical Physics
Randall Laboratory, Department of Physics, University of Michigan
Ann Arbor, MI 48109–1120, USA

Abstract. In M-theory vacua with vanishing 4-form $F_{(4)}$, one can invoke the ordinary Riemannian holonomy $H \subset SO(10,1)$ to account for unbroken supersymmetries $n = 1$, 2, 3, 4, 6, 8, 16, 32. However, the generalized holonomy conjecture, valid for non-zero $F_{(4)}$, can account for more exotic fractions of supersymmetry, in particular $16 < n < 32$. The conjectured holonomies are given by $\mathcal{H} \subset \mathcal{G}$ where $\mathcal{G}$ are the generalized structure groups $\mathcal{G} = SO(d-1,1) \times G(spacelike)$, $\mathcal{G} = ISO(d-1) \times G(null)$ and $\mathcal{G} = SO(d) \times G(timelike)$ with $1 \leq d < 11$. For example, $G(spacelike) = SO(16)$, $G(null) = [SU(8) \times U(1)] \ltimes \mathbb{R}^{56}$ and $G(timelike) = SO^*(16)$ when $d = 3$. Although extending spacetime symmetries, there is no conflict with the Coleman-Mandula theorem. The holonomy conjecture rules out certain vacua which are otherwise permitted by the supersymmetry algebra.

1 Introduction

M-theory not only provides a non-perturbative unification of the five consistent superstring theories, but also embraces earlier work on supermembranes and eleven-dimensional supergravity [1]. It is regarded by many as the dreamed-of final theory and has accordingly received an enormous amount of attention. It is curious, therefore, that two of the most basic questions of M-theory have until now remained unanswered:

i) *What are the symmetries of M-theory?*

ii) *How many supersymmetries can vacua of M-theory preserve?*

The first purpose of this paper is to argue that M-theory possesses previously unidentified hidden spacetime (timelike and null) symmetries in addition to the well-known hidden internal (spacelike) symmetries. These take the form of generalized structure groups $\mathcal{G}$ that replace the Lorentz group $SO(10,1)$.

The second purpose is to argue that the number of supersymmetries preserved by an M-theory vacuum is given by the number of singlets appearing in the decomposition of the 32-dimensional representation of $\mathcal{G}$ under $\mathcal{G} \supset \mathcal{H}$ where $\mathcal{H}$ are generalized holonomy groups.

⋆ MCTP-03-11 and hep-th/0303140

⋆⋆ Research supported in part by DOE Grant DE-FG02-95ER40899.

⋆⋆⋆ mduff@umich.edu

† jimliu@umich.edu

The equations of M-theory display the maximum number of supersymmetries N=32, and so n, the number of supersymmetries preserved by a particular vacuum, must be some integer between 0 and 32. But are some values of n forbidden and, if so, which ones? For quite some time it was widely believed that, aside from the maximal $n = 32$, n is restricted to $0 \leq n \leq 16$ with $n = 16$ being realized by the fundamental BPS objects of M-theory: the M2-brane, the M5-brane, the M-wave and the M-monopole. The subsequent discovery of intersecting brane configurations with $n = 0, 1, 2, 3, 4, 5, 6, 8, 16$ lent credence to this argument. In [2], on the other hand, it was shown that all values $0 \leq n \leq 32$ are allowed by the M-theory algebra [3], and examples of vacua with $16 < n < 32$ have indeed since been found. Following [4] and [5], we here put forward a *generalized holonomy conjecture* according to which the answer lies somewhere in between. Evidence in favor of this conjecture includes the observations that there are no known counterexamples and that a previously undiscovered example predicted in [5], namely n=14, has recently been found [6].

As we shall see, these conjectures are based on a group-theoretical argument which applies to the fully-fledged M-theory. To get the ball rolling, however, we begin with the low energy limit of M-theory, namely $D = 11$ supergravity. The unique $D = 11$ supermultiplet is comprised of a graviton g_{MN}, a gravitino Ψ_M and 3-form gauge field A_{MNP}, where $M = 0, 1, \ldots 10$, with 44, 128 and 84 physical degrees of freedom, respectively. In section 2, we conjecture that the supergravity equations of motion for this set of fields admit hidden timelike and null symmetries (in addition to previously demonstrated hidden spacelike ones). Then in section 3 we propose that, so long as the $D = 11$ Killing spinor equation has such hidden symmetries, we may enlarge the tangent space group into a generalized structure group. This allows us to analyze the number of supersymmetries based on a generalized holonomy conjecture. Partial justification for this conjecture is presented in section 4 in the context of a dimensionally reduced theory. In section 5 we discuss some consequences of generalized holonomy for classifying supersymmetric vacua, and finally conclude in section 6.

2 Hidden spacetime symmetries of D=11 supergravity

Long ago, Cremmer and Julia [7] pointed out that, when dimensionally reduced to d dimensions, $D = 11$ supergravity exhibits hidden symmetries. For example $E_7(global) \times SU(8)(local)$ when $d = 4$ and $E_8(global) \times SO(16)(local)$ when $d = 3$. The question was then posed [8]: do these symmetries appear magically only after dimensional reduction, or were they already present in the full uncompactified and untruncated $D = 11$ theory? The question was answered by de Wit and Nicolai [9,10] who made a $d/(11 - d)$ split and fixed the gauge by setting to zero the off-diagonal components of the elfbein. They showed that in the resulting field equations the local symmetries are indeed already present, but the global symmetries are not. For example, after making the split $SO(10, 1) \supset SO(3, 1) \times SO(7)$, we find the enlarged symmetry $SO(3, 1) \times SU(8)$. There is no global E_7 invariance (although the 70 internal components of the metric and 3-form may nevertheless be assigned to an $E_7/SU(8)$ coset). Similar results were

$d/(11-d)$	$\mathcal{G} = SO(d-1,1) \times G(spacelike)$	ϵ representation
10/1	$SO(9,1) \times \{1\}$	$16 + \overline{16}$
9/2	$SO(8,1) \times SO(2)$	$16_{\pm 1/2}$
8/3	$SO(7,1) \times SO(3) \times SO(2)$	$(8_s, 2)_{1/2} + (8_c, 2)_{-1/2}$
7/4	$SO(6,1) \times SO(5)$	$(8,4)$
6/5	$SO(5,1) \times SO(5) \times SO(5)$	$(4,4,1) + (\overline{4},1,4)$
5/6	$SO(4,1) \times USp(8)$	$(4,8)$
4/7	$SO(3,1) \times SU(8)$	$(2,1,8) + (1,2,\overline{8})$
3/8	$SO(2,1) \times SO(16)$	$(2,16)$
2/9	$SO(1,1) \times SO(16) \times SO(16)$	$(16,1)_{1/2} + (1,16)_{-1/2}$
1/10	$\{1\} \times SO(32)$	32

Table 1. Generalized structure groups: spacelike case. The last column denotes the representation of ϵ under $\mathcal{G}$.

found for other values of d: in each case the internal subgroup $SO(11-d)$ gets enlarged to some compact group $G(spacelike)$ while the spacetime subgroup $SO(d-1,1)$ remains intact[1]. In this paper we ask instead whether there are hidden *spacetime* symmetries. This is a question that could have been asked long ago, but we suspect that people may have been inhibited by the Coleman-Mandula theorem which forbids combining spacetime and internal symmetries [11]. However, this is a statement about Poincare symmetries of the S-matrix and here we are concerned with Lorentz symmetries of the equations of motion, so there will be no conflict.

The explicit demonstration of $G(spacelike)$ invariance by de Wit and Nicolai is very involved, to say the least. However, the result is quite simple: one finds the same $G(spacelike)$ in the full uncompactified $D = 11$ theory as was already found in the spacelike dimensional reduction of Cremmer and Julia. Here we content ourselves with the educated guess that the same logic applies to $G(timelike)$ and $G(null)$: they are the same as what one finds by timelike and null reduction, respectively. So we propose that, after making a $d/(11-d)$ split, the Lorentz subgroup $G = SO(d-1,1) \times SO(11-d)$ can be enlarged to the generalized structure groups $\mathcal{G} = SO(d-1,1) \times G(spacelike)$, $\mathcal{G} = ISO(d-1) \times G(null)$ and $\mathcal{G} = SO(d) \times G(timelike)$ as shown in Tables 1, 2 and 3.

Some of the noncompact groups appearing in the Tables may be unfamiliar, but a nice discussion of their properties may be found in [12]. For $d > 2$ the groups $G(spacelike)$, $G(timelike)$ and $G(null)$ are the same as those obtained from the spacelike dimensional reductions of Cremmer and Julia [7], the timelike reductions of Hull and Julia [13][2], and the null reduction of section 3.2, respectively. For our purposes, however, their physical interpretation is very different. They are here proposed as symmetries of the full $D = 11$ equations of motion; there is no compactification involved, whether toroidal or otherwise. This conjecture that these symmetries are present in the full theory and not merely in its

[1] We keep the terminology "spacetime" and "internal" even though no compactification or dimensional reduction is implied.

[2] Actually, for the 8/3 split, we have the factor $SO(1,1)$ instead of their $SO(2)$.

$d/(11-d)$	$\mathcal{G} = \mathrm{ISO}(d-1) \times G(null)$	ϵ representation
10/1	$\mathrm{ISO}(9)$	$16 + 16$
9/2	$\mathrm{ISO}(8) \times \mathbb{R}$	$8_s + 8_s + 8_c + 8_c$
8/3	$\mathrm{ISO}(7) \times \mathrm{ISO}(2) \times \mathbb{R}$	$8_{\pm 1/2} + 8_{\pm 1/2}$
7/4	$\mathrm{ISO}(6) \times [\mathrm{SO}(3) \times \mathrm{SO}(2)] \ltimes \mathbb{R}^6_{(3,2)}$	$(4,2)_{\pm 1/2} + (\overline{4},2)_{\pm 1/2}$
6/5	$\mathrm{ISO}(5) \times \mathrm{SO}(5) \ltimes \mathbb{R}^{10}_{(10)}$	$(4,4) + (4,4)$
5/6	$\mathrm{ISO}(4) \times [\mathrm{SO}(5) \times \mathrm{SO}(5)] \ltimes \mathbb{R}^{16}_{(4,4)}$	$(2,1,4,1) + (2,1,1,4)$ $+ (1,2,4,1) + (1,2,1,4)$
4/7	$\mathrm{ISO}(3) \times \mathrm{USp}(8) \ltimes \mathbb{R}^{27}_{(27)}$	$(2,8) + (2,8)$
3/8	$\mathrm{ISO}(2) \times [\mathrm{SU}(8) \times \mathrm{U}(1)] \ltimes \mathbb{R}^{56}_{(28_{1/2},\overline{28}_{-1/2})}$	$(8_{1/2})_{\pm 1/2} + (\overline{8}_{-1/2})_{\pm 1/2}$
2/9	$\mathbb{R} \times \mathrm{SO}(16) \ltimes \mathbb{R}^{120}_{(120)}$	$16 + 16$
1/10	$\{1\} \times [\mathrm{SO}(16) \times \mathrm{SO}(16)] \ltimes \mathbb{R}^{256}_{(16,16)}$	$(16,1) + (1,16)$

Table 2. Generalized structure groups: null case. The last column denotes the representation of ϵ under the maximum compact subgroup of $\mathcal{G}$.

$d/(11-d)$	$\mathcal{G} = \mathrm{SO}(d) \times G(timelike)$	ϵ representation
10/1	$\mathrm{SO}(10) \times \{1\}$	$16 + \overline{16}$
9/2	$\mathrm{SO}(9) \times \mathrm{SO}(1,1)$	$16_{\pm 1/2}$
8/3	$\mathrm{SO}(8) \times \mathrm{SO}(2,1) \times \mathrm{SO}(1,1)$	$(8_s,2)_{1/2} + (8_c,2)_{-1/2}$
7/4	$\mathrm{SO}(7) \times \mathrm{SO}(3,2)$	$(8,4)$
6/5	$\mathrm{SO}(6) \times \mathrm{SO}(5,\mathbb{C})$	$(4,4) + (\overline{4},\overline{4})$
5/6	$\mathrm{SO}(5) \times \mathrm{USp}(4,4)$	$(4,8)$
4/7	$\mathrm{SO}(4) \times \mathrm{SU}^*(8)$	$(2,1,8) + (1,2,\overline{8})$
3/8	$\mathrm{SO}(3) \times \mathrm{SO}^*(16)$	$(2,16)$
2/9	$\mathrm{SO}(2) \times \mathrm{SO}(16,\mathbb{C})$	$16_{1/2} + \overline{16}_{-1/2}$
1/10	$\{1\} \times \mathrm{SO}(16,16)$	32

Table 3. Generalized structure groups: timelike case. The last column denotes the representation of ϵ under $\mathcal{G}$.

dimensional reductions may be put to the test, however, as we shall later describe. For $d \leq 2$ it is less clear whether these generalized structure groups are actually hidden symmetries. See the caveats of section 4. The $\mathrm{SO}(16) \times \mathrm{SO}(16)$ for $d = 2$ is also discussed by Nicolai [14].

3 Hidden Symmetries and Generalized Holonomy

We begin by reviewing the connection between holonomy and the number of preserved supersymmetries, n, of supergravity vacua. This also serves to define our notation. Subsequently, we introduce a generalized holonomy which involves the hidden symmetries conjectured in the previous section.

3.1 Riemannian Holonomy

We are interested in solutions of the bosonic field equations

$$R_{MN} = \frac{1}{12}\left(F_{MPQR}F_N{}^{PQR} - \frac{1}{12}g_{MN}F^{PQRS}F_{PQRS}\right) \tag{1}$$

and

$$d * F_{(4)} + \frac{1}{2}F_{(4)} \wedge F_{(4)} = 0, \tag{2}$$

where $F_{(4)} = dA_{(3)}$. The supersymmetry transformation rule of the gravitino reduces in a purely bosonic background to

$$\delta\Psi_M = \tilde{D}_M \epsilon, \tag{3}$$

where the parameter ϵ is a 32-component anticommuting spinor, and where

$$\tilde{D}_M = D_M - \frac{1}{288}(\Gamma_M{}^{NPQR} - 8\delta^N_M \Gamma^{PQR})F_{NPQR}, \tag{4}$$

where Γ^A are the $D = 11$ Dirac matrices. Here D_M is the usual Riemannian covariant derivative involving the connection ω_M of the usual structure group $Spin(10, 1)$, the double cover of $SO(10, 1)$,

$$D_M = \partial_M + \frac{1}{4}\omega_M{}^{AB}\Gamma_{AB}. \tag{5}$$

The number of supersymmetries preserved by an M-theory background depends on the number of covariantly constant spinors,

$$\tilde{D}_M \epsilon = 0, \tag{6}$$

called *Killing* spinors. It is the presence of the terms involving the 4-form $F_{(4)}$ in (4) that makes this counting difficult. So let us first examine the simpler vacua for which $F_{(4)}$ vanishes. Killing spinors then satisfy the integrability condition

$$[D_M, D_N]\epsilon = \frac{1}{4}R_{MN}{}^{AB}\Gamma_{AB}\epsilon = 0, \tag{7}$$

where $R_{MN}{}^{AB}$ is the Riemann tensor. The subgroup of $Spin(10, 1)$ generated by this linear combination of $Spin(10, 1)$ generators Γ_{AB} corresponds to the *holonomy* group H of the connection ω_M. The number of supersymmetries, n, is then given by the number of singlets appearing in the decomposition of the 32 of $Spin(10, 1)$ under H. In Euclidean signature, connections satisfying (7) are automatically Ricci-flat and hence solve field equations when $F_{(4)} = 0$. In Lorentzian signature, however, they need only be Ricci-null [15] so Ricci-flatness has to be imposed as an extra condition. In Euclidean signature, the holonomy groups have been classified [16]. In Lorentzian signature, much less is known but the question of which subgroups H of $Spin(10, 1)$ leave a spinor invariant has been answered in [17]. There are two sequences according as the vector $v_A = \bar{\epsilon}\,\Gamma_A\,\epsilon$ is timelike or null, as shown in Tables 4 and 5. Since $v^2 \leq 0$, the spacelike v_A case does not arise. The

$d/(11-d)$	$H \subset SO(11-d) \subset Spin(10)$	n
7/4	$SU(2) \cong Sp(2)$	16
5/6	$SU(3)$	8
4/7	G_2	4
3/8	$SU(2) \times SU(2)$	8
	$Sp(4)$	6
	$SU(4)$	4
	$Spin(7)$	2
1/10	$SU(2) \times SU(3)$	4
	$SU(5)$	2

Table 4. Holonomy of static M-theory vacua with $F_{(4)} = 0$ and their supersymmetries.

$d/(11-d)$	$H \subset ISO(d-1) \times ISO(10-d) \subset Spin(10,1)$	n
10/1	$\mathbb{R}^9$	16
6/5	$\mathbb{R}^5 \times (SU(2) \ltimes \mathbb{R}^4)$	8
4/7	$\mathbb{R}^3 \times (SU(3) \ltimes \mathbb{R}^6)$	4
3/8	$\mathbb{R}^2 \times (G_2 \ltimes \mathbb{R}^7)$	2
2/9	$\mathbb{R} \times (SU(2) \ltimes \mathbb{R}^4) \times (SU(2) \ltimes \mathbb{R}^4)$	4
	$\mathbb{R} \times (Sp(4) \ltimes \mathbb{R}^8)$	3
	$\mathbb{R} \times (SU(4) \ltimes \mathbb{R}^8)$	2
	$\mathbb{R} \times (Spin(7) \ltimes \mathbb{R}^8)$	1

Table 5. Holonomy of non-static M-theory vacua with $F_{(4)} = 0$ and their supersymmetries.

timelike v_A case corresponds to static vacua, where $H \subset Spin(10) \subset Spin(10,1)$ while the null case to non-static vacua where $H \subset ISO(9) \subset Spin(10,1)$. It is then possible to determine the possible n-values [18,19] and one finds $n = 2, 4, 6, 8, 16, 32$ for static vacua, as shown in Table 4, and $n = 1, 2, 3, 4, 8, 16, 32$ for non-static vacua, as shown in Table 5.

3.2 Generalized holonomy

When we want to include vacua with $F_{(4)} \neq 0$ we face the problem that the connection in (4) is no longer the spin connection to which the bulk of the mathematical literature on holonomy groups is devoted. In addition to the $Spin(10,1)$ generators Γ_{AB}, it is apparent from (4) that there are terms involving Γ_{ABC} and Γ_{ABCDE}. As a result, the connection takes its values in the full $D = 11$ Clifford algebra. Moreover, this connection can preserve exotic fractions of supersymmetry forbidden by the Riemannian connection. For example, the M-branes at angles in [20] include n=5, the 11-dimensional pp-waves in [21,22,23,24] include $n = 18$, 20, 22, 24, 26 (and $n = 28$ for Type IIB), the squashed $N(1,1)$ spaces in [25] and the M5-branes in a pp-wave background in [26] include n=12 and the Gödel universes in [27] include $n = 18, 20, 22, 24$.

However, we can attempt to quantify this in terms of generalized holonomy groups $\mathcal{H} \subset \mathcal{G}$ where $\mathcal{G}$ are the generalized structure groups discussed in section 2. The generalized holonomy conjecture [4,5] states that one can assign a

holonomy $\mathcal{H} \subset \mathcal{G}$ to the generalized connection[3] appearing in the supercovariant derivative (4). Here we propose that, after making a $d/(11-d)$ split, the Lorentz subgroup $G = SO(d-1,1) \times SO(11-d)$ can be enlarged to the generalized structure groups $\mathcal{G} = SO(d-1,1) \times G(spacelike)$, $\mathcal{G} = ISO(d-1) \times G(null)$ and $\mathcal{G} = SO(d) \times G(timelike)$ as shown in Tables 1, 2 and 3. Note that in the right hand column of the tables we have listed the corresponding $\mathcal{G}$ representations under which the 32 supersymmetry parameters ϵ transform. The number of supersymmetries preserved by an M-theory vacuum is then given by the number of singlets appearing in the decomposition of these representations under $\mathcal{G} \supset \mathcal{H}$.

4 Structure groups from dimensional reduction

In this section we provide partial justification for the conjectured hidden symmetries by demonstrating their presence in the gravitino variation of the dimensionally reduced theory. In particular, we consider a spacelike dimensional reduction corresponding to a $d/(11-d)$ split. Turning on only d-dimensional scalars, the reduction ansatz is particularly simple

$$g^{(11)}_{MN} = \begin{pmatrix} \Delta^{-1/(d-2)} g_{\mu\nu} & 0 \\ 0 & g_{ij} \end{pmatrix}, \qquad A^{(11)}_{ijk} = \phi_{ijk}, \tag{8}$$

where $\Delta = \det g_{ij}$. For $d \leq 5$, we must also consider the possibility dualizing either $F_{(4)}$ components or (for $d = 3$) Kaluza-Klein vectors to scalars. We will return to such possibilities below. But for now we focus on $d \geq 6$. In this case, a standard dimensional reduction of the $D = 11$ gravitino transformation, (3), yields the d-dimensional gravitino transformation

$$\delta\psi_\mu = [D_\mu + \tfrac{1}{4} Q_\mu{}^{ab}\Gamma_{ab} + \tfrac{1}{24}\partial_\mu\phi_{ijk}\Gamma^{ijk}]\epsilon. \tag{9}$$

For completeness, we also note that the d-dimensional dilatinos transform according to

$$\delta\lambda_i = -\tfrac{1}{2}\gamma^\mu[P_{\mu\, ij}\Gamma^j - \tfrac{1}{36}(\Gamma_i{}^{jkl} - 6\delta^j_i\Gamma^{kl})\partial_\mu\phi_{jkl}]\epsilon. \tag{10}$$

In the above, the lower dimensional quantities are related to their $D = 11$ counterparts through

$$\begin{aligned} &\psi_\mu = \Delta^{\frac{1}{4(d-2)}}\left(\Psi^{(11)}_\mu + \frac{1}{d-2}\gamma_\mu\Gamma^i\Psi^{(11)}_i\right), \qquad \lambda_i = \Delta^{\frac{1}{4(d-2)}}\Psi^{(11)}_i, \\ &\epsilon = \Delta^{\frac{1}{4(d-2)}}\epsilon^{(11)}, \\ &Q^{ab}_\mu = e^{i[a}\partial_\mu e_i{}^{b]}, \qquad P_{\mu\, ij} = e^a_{(i}\partial_\mu e_{j)\, a}. \end{aligned} \tag{11}$$

We now see that the lower dimensional gravitino transformation, (9), may be written in terms of a covariant derivative under a generalized connection

$$\delta\psi_\mu = \hat{D}_\mu\epsilon, \qquad \hat{D}_\mu = \partial_\mu + \tfrac{1}{4}\Omega_\mu, \tag{12}$$

[3] A related conjecture was made in [28], where the generalized holonomy could be any subgroup of $SO(16,16)$. This also appears in our conjectured hidden structure groups under the $1/10$ split, though only in the timelike case $\mathcal{G}(timelike)$.

where

$$\Omega_\mu = \omega_\mu{}^{\alpha\beta}\gamma_{\alpha\beta} + Q_\mu{}^{ab}\Gamma_{ab} + \tfrac{1}{3!}e^{ia}e^{jb}e^{kc}\partial_\mu\phi_{ijk}\Gamma_{abc}. \tag{13}$$

Here γ_α are $SO(d-1,1)$ Dirac matrices, while Γ_a are $SO(11-d)$ Dirac matrices. This decomposition is suggestive of a generalized structure group with connection given by Ω_μ. However one additional requirement is necessary before declaring this an enlargement of $SO(d-1,1)\times SO(11-d)$, and that is to ensure that the algebra generated by Γ_{ab} and Γ_{abc} closes within itself. Along this line, we note that the commutators of these internal Dirac matrices have the schematic structure

$$[\Gamma^{(2)},\Gamma^{(2)}] = \Gamma^{(2)}, \qquad [\Gamma^{(2)},\Gamma^{(3)}] = \Gamma^{(3)}, \qquad [\Gamma^{(3)},\Gamma^{(3)}] = \Gamma^{(6)} + \Gamma^{(2)}. \tag{14}$$

Here the notation $\Gamma^{(n)}$ indicates the antisymmetric product of n Dirac matrices, and the right hand sides of the commutators only indicate what possible terms may show up. The first commutator above merely indicates that the Γ_{ab} matrices provide a representation of the Riemannian $SO(11-d)$ structure group.

For $d \geq 6$, the internal space is restricted to five or fewer dimensions. In this case, the antisymmetric product $\Gamma^{(6)}$ cannot show up, and the algebra clearly closes on $\Gamma^{(2)}$ and $\Gamma^{(3)}$. Working out the extended structure groups for these cases results in the expected Cremmer and Julia groups listed in the first four lines of Table 1. A similar analysis follows for $d \leq 5$. However, in this case, we must also dualize an additional set of fields to see the hidden symmetries. For $d = 5$, an additional scalar arises from the dual of $F_{\mu\nu\rho\sigma}$; this yields an addition to (13) of the form $\Omega_\mu^{\text{additional}} = \frac{1}{4!}\epsilon_\mu{}^{\nu\rho\sigma\lambda}F_{\mu\nu\rho\sigma}\Gamma_{123456}$. This $\Gamma^{(6)}$ term is precisely what is necessary for the closure of the algebra of (14). Of course, in this case, we must also make note of the additional commutators

$$[\Gamma^{(2)},\Gamma^{(6)}] = \Gamma^{(6)}, \qquad [\Gamma^{(3)},\Gamma^{(6)}] = \Gamma^{(7)}+\Gamma^{(3)}, \qquad [\Gamma^{(6)},\Gamma^{(6)}] = \Gamma^{(10)}+\Gamma^{(6)}+\Gamma^{(2)}. \tag{15}$$

However neither $\Gamma^{(7)}$ nor $\Gamma^{(10)}$ may show up in $d = 5$ for dimensional reasons.

The analysis for $d = 4$ is similar; however here

$$\Omega_\mu^{\text{additional}} = \frac{1}{3!}\epsilon_\mu{}^{\nu\rho\sigma}e^{ia}F_{\nu\rho\sigma i}\Gamma_a\Gamma_{1234567}.$$

Closure of the algebra on $\Gamma^{(2)}$, $\Gamma^{(3)}$ and $\Gamma^{(6)}$ then follows because, while $\Gamma^{(7)}$ may in principle arise in the middle commutator of (15), it turns out to be kinematically forbidden. For $d = 3$, on the other hand, in additional to a contribution $\Omega_\mu^{\text{additional}} = \frac{1}{2!\cdot 2!}\epsilon_\mu{}^{\nu\rho}e^{ia}e^{jb}F_{\nu\rho ij}\Gamma_{ab}\Gamma_{12345678}$, one must also dualize the Kaluza-Klein vectors $g_\mu{}^i$. Doing so gives rise to a $\Gamma^{(7)}$ in the generalized connection which, in addition to the previously identified terms, completes the internal structure group to $SO(16)$.

The remaining two cases, namely $d = 2$ and $d = 1$, fall somewhat outside the framework presented above. This is because in these low dimensions the generalized connections Ω_μ derived via reduction are partially incomplete. For $d = 2$, we find

$$\Omega_\mu^{(d=2)} = \omega_\mu{}^{\alpha\beta}\gamma_{\alpha\beta} + Q_\mu{}^{ab}\Gamma_{ab} + \tfrac{1}{9}(\delta_\mu^\nu - \tfrac{1}{2}\gamma_\mu{}^\nu)e^{ia}e^{jb}e^{kc}\partial_\nu\phi_{ijk}\Gamma_{abc}, \tag{16}$$

where $\gamma_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu}(\epsilon^{\alpha\beta}\gamma_{\alpha\beta})$ is necessarily proportional to the two-dimensional chirality matrix. Hence from a two-dimensional point of view, the scalars from the metric enter non-chirally, while the scalars from $F_{(4)}$ enter chirally. Taken together, the generalized connection (16) takes values in $SO(16)_+ \times SO(16)_-$, which we regard as the enlarged structure group. However not all generators are present because of lack of chirality in the term proportional to $Q_\mu{}^{ab}$. Thus at this point the generalized structure group deviates from the hidden symmetry group, which would be an infinite dimensional subgroup of affine E_8. Similarly, for $d = 1$, closure of the connection $\Omega_\mu^{(d=1)}$ results in an enlarged SO(32) structure group. However this is not obviously related to any actual hidden symmetry of the $1/10$ split.

Until now, we have considered the spacelike reductions leading to the generalized structure groups of Table 1. For a timelike reduction, we simply interchange a time and a space direction in the above analysis[4]. This results in an internal Clifford algebra with signature $(10 - d, 1)$, and yields the extended symmetry groups indicated in Table 3. Turning finally to the null case, we may replace one of the internal Dirac matrices with Γ_+ (where $+, -$ denote light-cone directions). Since $(\Gamma_+)^2 = 0$, this indicates that the extended structure groups for the null case are contractions the corresponding spacelike (or timelike) groups. In addition, by removing Γ_+ from the set of Dirac matrices, we essentially end up in the case of one fewer compactified dimensions. As a result, the $G(null)$ group in d-dimensions must have a semi-direct product structure involving the $G(spacelike)$ group in $(d + 1)$-dimensions. Of course, these groups also contain the original $ISO(10 - d)$ structure group as a subgroup. The resulting generalized structure groups are given in Table 2[5].

5 Counting supersymmetries

Having defined a generalized holonomy for vacua with $F_{(4)} \neq 0$, we now turn to some elementary examples. For the basic objects of M-theory, the M2-brane configuration may be placed under the 3/8 (spacelike) classification, as it has three longitudinal and eight transverse directions. Focusing on the transverse directions (which is the analog of looking at $\hat{D}_\mu$), the M2-brane has generalized holonomy SO(8) contained in $SO(2, 1) \times SO(16)$ [4]. In this case, the spinor decomposes as $(2, 16) = 2(8) + 16(1)$, indicating the expected presence of 16 singlets. For the M5-brane with 6/5 (spacelike) split, the generalized $\hat{D}_\mu$ holonomy is given by $SO(5)_+ \subset SO(5, 1) \times SO(5)_+ \times SO(5)_-$, with the spinor decomposition $(4, 4, 1) + (\bar{4}, 1, 4) = 4(4) + 16(1)$. Since the wave solution depends on nine space-like coordinates, we may regard it as a 1/10 (null) split. In this case, it has generalized $\tilde{D}_M$ holonomy $\mathbb{R}^9 \subset [SO(16) \times SO(16)] \ltimes \mathbb{R}^{256}_{(16,16)}$. The spinor again

[4] By postulating that the generalized structure groups survive as hidden symmetries of the full uncompactified theory, we avoid the undesirable features associated with compactifications including a timelike direction such as closed timelike curves.

[5] The reduction of D-dimensional pure gravity along a single null direction was analyzed by Julia and Nicolai [29].

decomposes into 16 singlets. Note, however, that since the wave is pure geometry, it could equally well be categorized under a 10/1 split as $\mathbb{R}^9 \subset \mathrm{ISO}(9)$. Finally, the KK monopole is described by a 7/4 (spacelike) split, and has $\hat{D}_\mu$ holonomy $\mathrm{SU}(2)_+ \subset \mathrm{SO}(6,1) \times \mathrm{SO}(5)$, where the spinor decomposes as $(8,4) = 8(2) + 16(1)$. In all four cases, the individual objects preserve exactly half of the 32 supersymmetries. However each object is associated with its own unique generalized holonomy, namely SO(8), SO(5), $\mathbb{R}^9$ and SU(2) for the M2, M5, MW and MK, respectively.

The supersymmetry of intersecting brane configurations may be understood in a similar manner based on generalized holonomy. For example, for a M5 and MK configuration sharing six longitudinal directions, we may choose a 6/5 split. In this case, the structure group is $\mathrm{SO}(5,1) \times \mathrm{SO}(5)_+ \times \mathrm{SO}(5)_-$, and the $\hat{D}_\mu$ holonomies of the individual objects are $\mathrm{SO}(5)_+$ and $\mathrm{SU}(2) \subset \mathrm{SO}(5)_{\mathrm{diag}}$, respectively. The holonomy for the combined configuration turns out to be $\mathrm{SO}(5)_+ \times \mathrm{SU}(2)_-$, with the spinor decomposing as $(4,4,1) + (\overline{4},1,4) = 4(4,1) + 4(1,2) + 8(1,1)$. The resulting eight singlets then signify the presence of a 1/4 supersymmetric configuration. In principle, this analysis may be applied to more general brane configurations. However one goal of understanding enlarged holonomy is to obtain a classification of allowed holonomy groups and, as a result, to obtain a unified treatment of counting supersymmetries. We now provide some observations along this direction.

We first note the elementary fact that a p-dimensional representation can decompose into any number of singlets between 0 and p, *except* $(p-1)$, since if we have $(p-1)$ singlets, we must have p. It follows that in theories with N supersymmetries, $n = N - 1$ is ruled out, even though it is permitted by the supersymmetry algebra.

In some cases, additional restrictions on n may be obtained. For example, if the supersymmetry charge transforms as the $(2,16)$ representation of $\mathcal{G}$ when $d = 3$, then n is restricted to 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 32 as first noted in [5]. No new values of n are generated by $d > 3$ reps. For example, the 4 of SO(5) can decompose only into 0, 2 or 4 singlets but not 1.

We note that all the even values of n discussed so far appear in the list and that $n = 30$ is absent. This is consistent with the presence of pp-waves with $n = 16, 18, 20, 22, 24, 26$ (and $n = 28$ for Type IIB) but the absence of n=30 noted in [24,21,22,23]. Of course a good conjecture should not only account for the existing data but should go on to predict something new. For example, Gell-Mann's flavor SU(3) not only accounted for the nine known members of the baryon decuplet but went on to predict the existence of the Ω^-, which was subsequently discovered experimentally. For M-theory supersymmetries, the role of the Ω^- is played by $n = 14$ which at the time of its prediction had not been discovered "experimentally". We note with satisfaction, therefore, that this missing member has recently been found in the form of a Gödel universe [6].

The $d = 2$ and $d = 1$ cases are more problematic since $\mathrm{SO}(16) \times \mathrm{SO}(16)$ and SO(32) in principle allow any n except $n = 31$. So more work is required to explain the presence of M-branes at angles with $n = 0, 1, 2, 3, 4, 5, 6, 8, 16$ but the absence of $n = 7$ noted in [20]. Presumably, a more detailed analysis will show

that only those subgroups compatible with these allowed values of n actually appear as generalized holonomy groups. The beginnings of a classification of all supersymmetric $D = 11$ solutions may be found in [30].

We can apply similar logic to theories with fewer than 32 supersymmetries. Of course, if M-theory really underlies all supersymmetric theories then the corresponding vacua will all be special cases of the above. However, it is sometimes useful to focus on such a sub-theory, for example the Type I and heterotic strings with $N = 16$. Here $G(spacelike) = SO(d) \times SO(d)$, $G(null) = ISO(d-1) \times ISO(d-1)$ and $G(timelike) = SO(d-1,1) \times SO(d-1,1)$. If the supersymmetry charge transforms as a $(2,8)$ representation of the generalized structure group when $d = 3$, then n is restricted to 0, 2, 4, 6, 8, 10, 12, 16. No new values of n are generated from other $d > 4$ reps. Once again, the $d = 2$ and $d = 1$ cases require a more detailed analysis.

6 The full M-theory

We have focused on the low energy limit of M-theory, but since the reasoning that led to the conjecture is based just on group theory, it seems reasonable to promote it to the full M-theory[6]. When counting the n value of a particular vacuum, however, we should be careful to note the phenomenon of *supersymmetry without supersymmetry*, where the supergravity approximation may fail to capture the full supersymmetry of an M-theory vacuum. For example, vacua related by T-duality and S-duality must, by definition, have the same n values. Yet they can appear to be different in supergravity [33,34], if one fails to take into account winding modes and non-perturbative solitons. So more work is needed to verify that the n values found so far in $D = 11$ supergravity exhaust those of M-theory, and to prove or disprove the conjecture.

Notes added

After this paper was posted on the archive, a very interesting paper by Hull appeared [35] which generalizes and extends the present theme. Hull conjectures that the hidden symmetry of M-theory is as large as $SL(32,\mathbb{R})$ and that this is necessary in order to accommodate all possible generalized holonomy groups. We here make some remarks in the light of Hull's paper:

Hidden symmetries:

Hull stresses that, as a candidate hidden symmetry, $SL(32,\mathbb{R})$ is background independent. However, the hidden symmetries displayed in Tables 1, 2 and 3 are also background independent. They depend only on the choice of non-covariant split and gauge in which to write the field equations. Hull's proposal is nevertheless very attractive since $SL(32,\mathbb{R})$ contains all the groups in Tables 1, 2 and 3 as subgroups and would thus answer the question of whether all these symmetries are present at the same time.

[6] Similar conjectures can be applied to M-theory in signatures (9,2) and (6,5) [31], the so-called M' and M^* theories [32], but the groups will be different.

One can accommodate $SL(32,\mathbb{R})$ by extending the $d/(11-d)$ split to include the $d=0$ case. Then the same $SL(32,\mathbb{R})$ would appear in all three tables. At the other end, one could also include the $d=11$ case. Then the same $SO(10,1)$ would appear in all three tables. Our reason for not including the $d=0$ case stems from the apparent need to make a non-covariant split and to make the corresponding gauge choice before the hidden symmetries become apparent [9,10]. Moreover, from the point of view of guessing the hidden symmetries from the dimensional reduction, the $d=0$ case would be subject to the same caveats as the $d=1$ and $d=2$ cases: not all group generators are present in the covariant derivative. $SL(32,\mathbb{R})$ requires $\{\Gamma^{(1)},\Gamma^{(2)},\Gamma^{(3)},\Gamma^{(4)},\Gamma^{(5)}\}$ whereas only $\{\Gamma^{(2)},\Gamma^{(3)},\Gamma^{(5)}\}$ appear in the covariant derivative. This is an important issue deserving of further study. That M-theory could involve a $GL(32,\mathbb{R})$ has also been conjectured by Barwald and West [36].

Generalized holonomy:

Hull goes on to stress the importance of $SL(32,\mathbb{R})$ by finding solutions whose holonomy is contained in $SL(32,\mathbb{R})$ but not in Tables 1, 2 and 3. Although not all generators are present in the covariant derivative, they are all present in the commutator. So we agree with Hull that $SL(32,\mathbb{R})$ is necessary if one wants to embrace all possible generalized holonomies.

Indeed, since the basic objects of M-theory discussed in section 5 involve warping by a harmonic function, the $\hat{D}$ holonomy is smaller than the $\tilde{D}$ holonomy, which requires extra $\mathbb{R}^n$ factors. Interestingly enough, the $\hat{D}$ holonomy nevertheless yields the correct counting of supersymmetries.

Hull points out that, in contrast to the groups appearing in Tables 1, 2 and 3, $SL(32,\mathbb{R})$ does not obey the $n\neq N-1$ rule of section 5, and hence M-theory vacua with $n=31$ are in principle possible[7]. Of course we do not yet know whether the required $\mathbb{R}^{31}$ holonomy actually appears. To settle the issue of which n values are allowed, it would be valuable to do for supergravity what Berger [16] did for gravity and have a complete classification of all possible generalized holonomy groups. But this may prove quite difficult.

So we remain open-minded about a formulation of M-theory with $SL(32,\mathbb{R})$ symmetry, but acknowledge the need for $SL(32,\mathbb{R})$ from the point of view of generalized holonomy.

Acknowledgments

We have enjoyed useful conversations with Hisham Sati.

References

1. For a review, see M. J. Duff, *The world in eleven dimensions: supergravity, supermembranes and M-theory* (I.O.P. Publishing, Bristol, 1999).

[7] The case for $n=31$ has also been made by Bandos *et al.* [37] in the different context of hypothetical preons of M-theory preserving 31 out of 32 supersymmetries.

2. J. P. Gauntlett and C. M. Hull, *BPS states with extra supersymmetry*, JHEP **0001**, 004 (2000) [hep-th/9909098].
3. P. K. Townsend, *M-theory from its superalgebra*, hep-th/9712004.
4. M. J. Duff and K. Stelle, *Multimembrane solutions of* $d = 11$ *supergravity*, Phys. Lett. B **253**, 113 (1991).
5. M. J. Duff, *M-theory on manifolds of* G_2 *holonomy: the first twenty years*, hep-th/0201062.
6. T. Harmark and T. Takayanagi, *Supersymmetric Gödel universes in string theory*, hep-th/0301206.
7. E. Cremmer and B. Julia, *The* $SO(8)$ *supergravity*, Nucl. Phys. B **159**, 141 (1979).
8. M. J. Duff, $E_8 \times SO(16)$ *Symmetry of* $d = 11$ *Supergravity?*, in *Quantum Field Theory and Quantum Statistics, Essays in Honor of E. S. Fradkin*, Batalin, Isham and Vilkovisky, eds. (Adam Hilger, 1986).
9. B. de Wit and H. Nicolai, $D = 11$ *supergravity with local* $SU(8)$ *invariance*, Nucl. Phys. B **274**, 363 (1986).
10. H. Nicolai, $D = 11$ *supergravity with local* $SO(16)$ *invariance*, Phys. Lett. B **187**, 316 (1987).
11. S. R. Coleman and J. Mandula, *All possible symmetries of the S-matrix*, Phys. Rev. **159**, 1251 (1967).
12. R. Gilmore, *Lie groups, Lie algebras and some of their applications* (Wiley, New York, 1974).
13. C. Hull and B. Julia, *Duality and moduli spaces for time-like reductions*, Nucl. Phys. B **534**, 250 (1998) [hep-th/9803239].
14. H. Nicolai, *Two-dimensional gravities and supergravities as integrable system*, in *Recent aspects of quantum fields: proceedings of the XXX Int. Universitätswochen für Kernphysik, Schladming, Austria, Feb 27 — Mar 5, 1991*, H. Mitter and H. Gausterer, eds. (Springer-Verlag, Berlin, 1991).
15. J. Figueroa-O'Farrill, *Breaking the M-waves*, Class. Quant. Grav. **17**, 2925 (2000) [hep-th/9904124].
16. M. Berger, *Sur les groupes d'holonomie homogene des varietes a connexion affine et des varietes riemanniennes*, Bull. Soc. Math. France **83** (1955) 225.
17. R. L. Bryant, *Pseudo-Riemannian metrics with parallel spinor fields and non-vanishing Ricci tensor*, math.DG/0004073.
18. B. S. Acharya, J. M. Figueroa-O'Farrill and B. Spence, *Planes, branes and automorphisms. I: Static branes*, JHEP **9807**, 004 (1998) [hep-th/9805073].
19. B. S. Acharya, J. M. Figueroa-O'Farrill, B. Spence and S. Stanciu, *Planes, branes and automorphisms. II: Branes in motion*, JHEP **9807**, 005 (1998) [hep-th/9805176].
20. N. Ohta and P. K. Townsend, *Supersymmetry of M-branes at angles* Phys. Lett. B **418**, 77 (1998) [hep-th/9710129].
21. J. Michelson, *(Twisted) toroidal compactification of pp-waves*, Phys. Rev. D **66**, 066002 (2002) [hep-th/0203140].
22. M. Cvetic, H. Lu and C. N. Pope, *M-theory pp-waves, Penrose limits and supernumerary supersymmetries*, Nucl. Phys. B **644**, 65 (2002) [hep-th/0203229].
23. J. P. Gauntlett and C. M. Hull, *pp-waves in 11-dimensions with extra supersymmetry*, JHEP **0206**, 013 (2002) [hep-th/0203255].
24. I. Bena and R. Roiban, *Supergravity pp-wave solutions with 28 and 24 supercharges*, hep-th/0206195.
25. D. N. Page and C. N. Pope, *New squashed solutions of* $d = 11$ *supergravity*, Phys. Lett. B **147**, 55 (1984).
26. H. Singh, *M5-branes with 3/8 supersymmetry in pp-wave background*, hep-th/0205020.

27. J. P. Gauntlett, J. B. Gutowski, C. M. Hull, S. Pakis and H. S. Reall, *All supersymmetric solutions of minimal supergravity in five dimensions*, hep-th/0209114.
28. M. Berkooz, M. Douglas and R. G. Leigh, *Branes intersecting at angles*, Nucl. Phys. B **480**, 265 (1996) [hep-th/9606139].
29. B. Julia and H. Nicolai, *Null Killing vector dimensional reduction and Galilean geometrodynamics*, Nucl. Phys. B **439**, 291 (1995) [hep-th/9412002].
30. J. P. Gauntlett and S. Pakis, *The geometry of* $D = 11$ *Killing spinors*, JHEP **0304**, 039 (2003) [hep-th/0212008].
31. M. P. Blencowe and M. J. Duff, *Supermembranes And The Signature Of Space-Time*, Nucl. Phys. B **310**, 387 (1988).
32. C. M. Hull, JHEP **9811**, 017 (1998) [hep-th/9807127].
33. M. J. Duff, H. Lü and C. N. Pope, *Supersymmetry without supersymmetry*, Phys. Lett. B **409**, 136 (1997) [hep-th/9704186].
34. M. J. Duff, H. Lü and C. N. Pope, $AdS_5 \times S^5$ *untwisted*, Nucl. Phys. B **532**, 181 (1998) [hep-th/9803061].
35. C. Hull, *Holonomy and symmetry in M-theory*, hep-th/0305039.
36. O. Barwald and P. C. West, *Brane rotating symmetries and the fivebrane equations of motion*, Phys. Lett. B **476**, 157 (2000) [hep-th/9912226].
37. I. A. Bandos, J. A. de Azcarraga, J. M. Izquierdo and J. Lukierski, *BPS states in M-theory and twistorial constituents*, Phys. Rev. Lett. **86**, 4451 (2001) [hep-th/0101113].

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 211)*
Portorož, Slovenia, July 12-17, 2003

On the Resolution of Space-Time Singularities II

Marco Maceda[1] and John Madore[1,2]

[1] Laboratoire de Physique Théorique
Université de Paris-Sud, Bâtiment 211, F-91405 Orsay
[2] Max-Planck-Institut für Physik
Föhringer Ring 6, D-80805 München

Abstract. In previous articles it has been argued that a differential calculus over a noncommutative algebra uniquely determines a gravitational field in the commutative limit and that there is a unique metric which remains as a commutative 'shadow'. Some examples were given of metrics which resulted from a given algebra and given differential calculus. Here we aboard the inverse problem, that of constructing the algebra and the differential calculus from the commutative metric. As an example a noncommutative version of the Kasner metric is proposed which is periodic.

1 Motivation

A definition has been given [1] of a torsion-free metric-compatible linear connection on a differential calculus $\Omega^*(\mathcal{A})$ over an algebra $\mathcal{A}$ which has certain rigidity properties provided that the center $\mathcal{Z}(\mathcal{A})$ of $\mathcal{A}$ is trivial. It was argued from simple examples that a differential calculus over a noncommutative algebra uniquely determines a gravitational field in the commutative limit. Some examples have been given [2,3,4] of metrics which resulted from a given algebra and given differential calculus. Here we aboard the inverse problem, that of constructing the algebra and the differential calculus from the commutative metric. As an example we construct noncommutative versions of the Kasner metric and we show that it is possible to choose an algebra such that the metric is nonsingular before taking the commutative limit. The 'II' on the title alludes to a preliminary version given at the Torino Euroconference [5] on noncommutative geometry [6].

The physical idea we have in mind is that the description of space-time using a set of commuting coordinates is only valid at length scales greater than some fundamental length. At smaller scales it is impossible to localize a point and a new geometry must be used. We can use a solid-state analogy and think of the ordinary Minkowski coordinates as macroscopic order parameters obtained by 'course-graining' over regions whose size is determined by a fundamental area scale $\kappa\!\!\!\!-$, which is presumably, but not necessarily, of the order of the Planck area $G\hbar$. They break down and must be replaced by elements of a noncommutative algebra when one considers phenomena on smaller scales. A simple visualization is afforded by the orientation order parameter of nematic liquid crystals. The commutative free energy is singular in the core region of a disclination. There is

of course no physical singularity; the core region can simply not be studied using the commutative order parameter.

As a concrete example we have chosen, for historical reasons, the Kasner metric; we show that its singularity can be resolved into an essentially noncommutative structure. We do not however claim that an arbitrary singularity in a metric on an arbitrary smooth manifold can be resolved using a noncommutative structure. From the point of view we are adopting a commutative geometry is a rather singular limit. The close relation between the differential calculus and the metric can at most be satisfied when the center is trivial. This manifests itself in the fact that on an ordinary manifold one can put any metric with any singularity. We argue only that those metrics which are 'physical' in some sense, for example are Ricci flat, can have resolvable singularities.

There is a similarity of the method we use to resolve the singularity with the method known in algebraic geometry as 'blowing up' a singularity [7] as well as with the method used by 't Hooft and Polyakov to resolve the monopole singularity. The regular solution found in this case can in fact be considered as the Dirac monopole solution on a noncommutative geometry which contains the 2×2 matrix algebra as extra factor. Since we are now dealing with a dynamical field configuration it is improbable that the singularity will admit being blown up by a finite-dimensional algebra. Our solutions offer evidence however in favor of this possibility with an algebra of infinite dimension.

In previous articles the algebra and the differential calculus were given and the linear connection and metric were constructed. It was argued [8,9] that given the algebra $\mathcal{A}$ the structure of $\Omega^*(\mathcal{A})$ is intimately connected with the gravitational field which remains on V as shadow in the commutative limit $\bar{k} \to 0$. Within the general framework which we here consider, the principal difference between the commutative and noncommutative cases lies in the spectrum of the operators which we use to generate the noncommutative algebra which replaces the algebra of functions. This in turn depends not only on the structure of this algebra as abstract algebra but on the representation of it which we choose to consider. Here we attempt the inverse problem, that of constructing the algebra *and* the differential calculus from the commutative linear connection. We cannot claim that the procedure is in any way unique.

For a discussion of the relation of noncommutative geometry to the problem of space-time singularities from rather different points of view from the one we adopt we refer, for example, to Heller & Sasin [10], to Hawkins [11] or to Lizzi *et al.* [12]. For a recent discussion of diffeomorphism invariance within the context of commutative gravity we refer to Gaul & Rovelli [13]. For a discription of a noncommutative approach to gravity using a choice of metric which does not fulfill the criteria which we use we refer to Aschieri & Castellani [14]. There seems to be a relation between the quadratic momentum algebra we use and non-linear representations of momentum operators considered recently. We refer to Chakrabarti [15] for review of this and references to the previous literature. We refer elsewhere for a description of the same 'quantization' applied to the PP wave [16] and for a possible cosmological application [17].

In the next sections we introduce the general formalism of noncommutative geometry which we use and we make some general remarks concerning the problem of 'quantization' of space-time. In Section 7 we recall the commutative Kasner metric. In Sections 8 and 9 we study the structure of the algebra we associate to the metric; we make also some remarks concerning perturbative approximations to noncommutative geometry and present the Kasner solution as a perturbative solution in $\bar{k}$. The key formulae are (52) and (53). They can be used to construct other examples without knowledge of the preceding material.

There is much that needs further work. We have not explicitly constructed the differential calculus nor have we examined in detail the complex structure of the algebra. There seems to be evidence of a cosmological constant associated to the noncommutativity but this remains elusive. All we can affirm is that a noncommutative structure is similar in certain aspects to extra dimensions and so could be expected to yield an effective cosmological constant in dimension four in the same way that extra (Ricci non-flat) dimensions give rise to one.

Greek indices take values from 0 to 3; the first half of the alphabet is used to index (moving) frames and the second half to index generators. Latin indices a, b, *etc.* take values from 1 to 3 and the indices i, j, *etc.* values from 0 to $n-1$.

2 The general formalism

The notation is the same as that of a previous article [5] on the symplectic structure of space-time and is based on a noncommutative generalization [18,8,19] of the Cartan moving-frame formalism . Let $\mathcal{A} = \mathcal{C}(V)$ be the algebra of smooth real-valued functions on a space-time V which for simplicity we shall suppose parallelizable and with a metric and linear connection defined in terms of a globally defined moving frame θ^α. Let $\Omega^*(\mathcal{A})$ be the algebra of de Rham differential forms. The space $\Omega^1(\mathcal{A})$ of 1-forms is free of rank 4 as a $\mathcal{A}$-module. According to the general idea outlined above a singularity in the metric is due to the use of commuting coordinates beyond their natural domain of definition into a region where they are physically inappropriate. From this point of view the space-time V should be more properly described 'near the singularity' by a noncommutative $*$-algebra $\mathcal{A}$ over the complex numbers with four hermitian generators x^λ. The observables will be some subset of the hermitian elements of $\mathcal{A}$. We shall not discuss this problem here; we shall implicitly suppose that all hermitian elements of $\mathcal{A}$ are observables, including the 'coordinates'. We shall not however have occasion to use explicitly this fact.

We introduce 6 additional elements $J^{\mu\nu}$ of $\mathcal{A}$ by the relations

$$[x^\mu, x^\nu] = i\bar{k} J^{\mu\nu}. \tag{1}$$

The details of the structure of $\mathcal{A}$ will be contained for example in the commutation relations $[x^\lambda, J^{\mu\nu}]$. One can define recursively an infinite sequence of elements by setting for $p \geq 1$

$$[x^\lambda, J^{\mu_1\cdots\mu_p}] = i\bar{k} J^{\lambda\mu_1\cdots\mu_p}. \tag{2}$$

We shall assume that for the description of a generic (strong) gravitational field the appropriate algebra $\mathcal{A}$ has a trivial center $\mathcal{Z}(\mathcal{A})$:

$$\mathcal{Z}(\mathcal{A}) = \mathbb{C}. \tag{3}$$

The only argument we have in favour of this assumption is the fact that it would be difficult to interpret the meaning of the center. The x^μ will be referred to as 'position generators'. We shall suppose also that there is a set of $n (= 4)$ antihermitian 'momentum generators' λ_α and a 'Fourier transform'

$$F : x^\mu \longrightarrow \lambda_\alpha = F_\alpha(x^\mu)$$

which takes the position generators to the momentum generators.

Let ρ be a representation of $\mathcal{A}$ as an algebra of linear operators on some Hilbert space. For every $k_\mu \in \mathbb{R}^4$ one can construct a unitary element $u(k) = e^{ik_\mu x^\mu}$ of $\mathcal{A}$ and one can consider the weakly closed algebra $\mathcal{A}_\rho$ generated by the image of the $u(k)$ under ρ. The momentum operators λ_α are also unbounded but using them one can construct also a set of 'translation' operators $\hat{u}(\xi) = e^{\xi^\alpha \lambda_\alpha}$ whose image under ρ belongs also to $\mathcal{A}_\rho$. In general $\hat{u}u \neq u\hat{u}$; if the metric which we introduce is the flat metric then we shall see that $[\lambda_\alpha, x^\mu] = \delta^\mu_\alpha$ and in this case we can write the commutation relations $\hat{u}u = qu\hat{u}$ with $q = e^{ik_\mu \xi^\mu}$; the 'Fourier transform' is the simple linear transformation

$$\lambda_\alpha = \frac{1}{i\kern-0.3em k} J^{-1}_{\alpha\mu} x^\mu \tag{4}$$

for some symplectic structure $J^{\alpha\mu}$. If the structure is degenerate then it is no longer evident that the algebra can be generated by either the position generators or the momentum generators alone. In such cases we define the algebra $\mathcal{A}$ to be the one generated by both sets. The derivations could be considered as outer derivations of the smaller algebra generated by the x^μ; they become inner in the extended algebra,

We shall suppose that $\mathcal{A}$ has a commutative limit which is an algebra $\mathcal{C}(V)$ of smooth functions on a space-time V endowed with a globally defined moving frame θ^α and thus a metric. By parallelizable we mean that the module $\Omega^1(\mathcal{A})$ has a basis θ^α which commutes with the elements of $\mathcal{A}$. For all $f \in \mathcal{A}$

$$f\theta^\alpha = \theta^\alpha f. \tag{5}$$

We shall see that this implies that the metric components must be constants, a condition usually imposed on a moving frame. It also means that we have 'frozen out' local Lorentz transformations since they do not leave this condition invariant. The frame θ^α allows one [20] to construct a representation of the differential algebra from that of $\mathcal{A}$. Following strictly what one does in ordinary geometry, we shall introduce the set of derivations e_α to be dual to the frame θ^α, that is with

$$\theta^\alpha(e_\beta) = \delta^\alpha_\beta. \tag{6}$$

We define the differential exactly as did E. Cartan in the commutative case. If e_α is a derivation of $\mathcal{A}$ then for every element $f \in \mathcal{A}$ we define df by the constraint

$df(e_\alpha) = e_\alpha f$. The differential calculus is defined as the largest one consistent with the module structure of the 1-forms so constructed. One can at this point take the classical limit to obtain four functions $\tilde{\lambda}_\alpha(\tilde{x}^\mu)$ which satisfy the equations

$$\{\tilde{\lambda}_\alpha, \tilde{x}^\mu\} = \tilde{e}^\mu_\alpha.$$

This defines a Poisson structure directly from which one can calculate the $\{\tilde{x}^\mu, \tilde{x}^\nu\}$. In this way only at the last moment does one pass to a noncommutative algebra and most of the problem remains within the category of smooth manifolds.

It follows from general arguments that the momenta λ_α must satisfy the consistency condition

$$2\lambda_\gamma\lambda_\delta P^{\gamma\delta}{}_{\alpha\beta} - \lambda_\gamma F^\gamma{}_{\alpha\beta} - K_{\alpha\beta} = 0. \tag{7}$$

The $P^{\gamma\delta}{}_{\alpha\beta}$ define the product π in the algebra of forms:

$$\theta^\alpha\theta^\beta = P^{\alpha\beta}{}_{\gamma\delta}\theta^\gamma \otimes \theta^\delta. \tag{8}$$

This product is defined to be the one with the least relations which is consistent with the module structure of the 1-forms. The $F^\gamma{}_{\alpha\beta}$ are related to the 2-form $d\theta^\alpha$ through the structure equations:

$$d\theta^\alpha = -\frac{1}{2}C^\alpha{}_{\beta\gamma}\theta^\beta\theta^\gamma. \tag{9}$$

In the noncommutative case the structure elements are defined as

$$C^\alpha{}_{\beta\gamma} = F^\alpha{}_{\beta\gamma} - 2\lambda_\delta P^{(\alpha\delta)}{}_{\beta\gamma}. \tag{10}$$

It follows that

$$e_\alpha C^\alpha{}_{\beta\gamma} = 0. \tag{11}$$

This must be imposed then at the classical level and can be used as a gauge-fixing condition. We impose also, without loss of generality the conditions

$$F^\eta{}_{\alpha\beta}P^{\alpha\beta}{}_{\gamma\delta} = F^\eta{}_{\gamma\delta}, \qquad K_{\alpha\beta}P^{\alpha\beta}{}_{\gamma\delta} = K_{\gamma\delta}.$$

There follows a similar relation for the $F^\gamma{}_{\alpha\beta}$.

Finally, to complete the definition of the coefficients of the consistency condition (7) we introduce the special 1-form $\theta = -\lambda_\alpha\theta^\alpha$. In the commutative, flat limit

$$\theta \to i\partial_\alpha dx^\alpha.$$

As an (antihermitian) 1-form θ defines a covariant derivative on an associated $\mathcal{A}$-module with local gauge transformations given by the unitary elements of $\mathcal{A}$. The $K_{\alpha\beta}$ are related to the curvature of θ:

$$d\theta + \theta^2 = K, \qquad K = -\frac{1}{2}K_{\alpha\beta}\theta^\alpha\theta^\beta.$$

All the coefficients lie in the center $\mathcal{Z}(\mathcal{A})$ of the algebra.

The condition (7) can be expressed also in terms of a twisted commutator

$$[\lambda_\alpha, \lambda_\beta]_P = 2P^{\gamma\delta}{}_{\alpha\beta}\lambda_\gamma\lambda_\delta$$

as

$$[\lambda_\alpha, \lambda_\beta]_P = \lambda_\gamma F^\gamma{}_{\alpha\beta} + K_{\alpha\beta}.$$

It is also connected with the condition that $d^2 f = 0$. The differential df of an element $f \in \mathcal{A}$ is given by $df = e_\alpha f \theta^\alpha$. Since, in particular

$$d^2\lambda_\gamma = d([\lambda_\beta, \lambda_\gamma]\theta^\beta) = ([\lambda_\alpha, [\lambda_\beta, \lambda_\gamma]] - \tfrac{1}{2}[\lambda_\mu, \lambda_\gamma]C^\mu{}_{\alpha\beta})\theta^\alpha\theta^\beta$$

it follows that

$$P^{\alpha\beta}{}_{\gamma\delta}e_\alpha e_\beta - C^\gamma{}_{\alpha\beta}e_\gamma = 0.$$

This is the same as Equation (7).

We must now compare in some way the commutators $[\lambda_\alpha, \lambda_\beta]$ and $[e_\alpha, e_\beta]$. Consider $C^\alpha{}_{\beta\gamma}$ as defined by the structure equation (9). Suppose further that for some numbers c_1 and c_2 the relations

$$[\lambda_\alpha, \lambda_\beta] = c_1 C^\gamma{}_{\alpha\beta}\lambda_\gamma, \qquad [e_\alpha, e_\beta] = c_2 C^\gamma{}_{\alpha\beta}e_\gamma$$

hold. The identities

$$([e_\alpha, e_\beta] - C^\gamma{}_{\alpha\beta}e_\gamma)x^\mu = (c_2 - c_1)C^\gamma{}_{\alpha\beta}e_\gamma x^\mu + [C^\gamma{}_{\alpha\beta}, x^\mu]\lambda_\gamma \tag{12}$$

place restrictions on the coefficients. In the commutative limit one must have $c_1 = c_2 = 1$. In a general noncommutative geometry there is no relation between the 2-form $d\theta^\alpha$ and the commutator $[e_\alpha, e_\beta]$. Such a relation would fix the value of c_2. In the formalism we are here considering the $C^\gamma{}_{\alpha\beta}$ are linear functionals of the momentum generators. (The theory has that is only four degrees of freedom and not the ten of general relativity. There are six which have been fixed by the choice of frame but they are not to be identified with the missing six.) It follows that

$$[C^\gamma{}_{\alpha\beta}, x^\mu]\lambda_\gamma = C^\gamma{}_{\alpha\beta}e_\gamma x^\mu.$$

Because of the Leibniz rule then one must have in general the relation $c_2 = 2c_1$. We shall choose $c_2 = 1$ so that the relation of the commutative limit is satisfied in general. We are forced then in general to choose $c_1 = \frac{1}{2}$. We shall assume that the commutator and the gravitational field are two aspects of the same phenomenon which we call gravity. We therefore suppose that in a 'realistic' situation both are present and one should take into account the relation found between c_1 and c_2.

3 The algebra

Equation (10) is the correspondence principle which associates a differential calculus to a metric. On the left in fact the quantity $C^\alpha{}_{\beta\gamma}$ determines a moving frame, which in turn fixes a metric; on the right are the elements of the algebra which fix to a large extent the differential calculus. A 'blurring' of a geometry proceeds via this correspondence. It is evident that in the presence of curvature the

1-forms cease to anticommute. On the other hand it is possible for flat 'space' to be described by 'coordinates' which do not commute. The correspondence principle between the classical and noncommutative geometries can be also described as the map

$$\tilde{\theta}^\alpha \mapsto \theta^\alpha \tag{13}$$

with the product satisfying the condition

$$\tilde{\theta}^\alpha \tilde{\theta}^\beta \mapsto P^{\alpha\beta}{}_{\gamma\delta}\theta^\gamma \theta^\delta.$$

The tilde on the left is to indicate that it is the classical form. The condition can be written also as

$$\tilde{C}^\alpha{}_{\beta\gamma} \mapsto C^\alpha{}_{\eta\zeta} P^{\eta\zeta}{}_{\beta\gamma}$$

or as

$$\lim_{k\to 0} C^\alpha{}_{\beta\gamma} = \tilde{C}^\alpha{}_{\beta\gamma}. \tag{14}$$

A solution to these equations would be a solution to the problem we have set. It would be however unsatisfactory in that no smoothness condition has been imposed. This can at best be done using the inner derivations. We shall construct therefore the set of momentum generators. The procedure we shall follow is not always valid; a counter example has been constructed [21] for the flat metric on the torus. The correspondence principle which in fact we shall actually use is a modified version of the map

$$\tilde{e}_\alpha \mapsto \lambda_\alpha$$

which is the inverse of that introduced by von Neumann to represent the Heisenberg algebra.

We introduce an involution [22] on the algebra of forms using [23] a reality condition on derivations, a procedure which is more or less a straightforward generalization of that which is used in the case of ordinary differential manifolds. The involution depends on the form of the product projection π. For general $\xi, \eta \in \Omega^1(\mathcal{A})$ it follows that

$$(\xi\eta)^* = -\eta^*\xi^*.$$

In particular

$$(\theta^\alpha\theta^\beta)^* = -\theta^\beta\theta^\alpha.$$

The product of two frame elements is hermitian then if and only if they anticommute. Recall that the product of two hermitian elements f and g of the algebra is hermitian if and only if they commute. When the frame exists one has necessarily also the relations

$$(f\xi\eta)^* = (\xi\eta)^* f^*, \qquad (f\xi \otimes \eta)^* = (\xi \otimes \eta)^* f^*$$

for arbitrary $f \in \mathcal{A}$.

We write $P^{\alpha\beta}{}_{\gamma\delta}$ in the form

$$P^{\alpha\beta}{}_{\gamma\delta} = \tfrac{1}{2}\delta^{[\alpha}_\gamma \delta^{\beta]}_\delta + i k Q^{\alpha\beta}{}_{\gamma\delta} \tag{15}$$

of a standard projector plus a perturbation and we decompose $Q^{\alpha\beta}{}_{\gamma\delta}$ as the sum of two terms

$$Q^{\alpha\beta}{}_{\gamma\delta} = Q_-^{\alpha\beta}{}_{\gamma\delta} + Q_+^{\alpha\beta}{}_{\gamma\delta}$$

symmetric (antisymmetric) and antisymmetric (symmetric) with respect to the upper (lower) indices. The condition that $P^{\alpha\beta}{}_{\gamma\delta}$ be a projector is satisfied to first order in k because of the property that

$$Q^{\alpha\beta}{}_{\gamma\delta} = P^{\alpha\beta}{}_{\zeta\eta} Q^{\zeta\eta}{}_{\gamma\delta} + Q^{\alpha\beta}{}_{\zeta\eta} P^{\zeta\eta}{}_{\gamma\delta}.$$

The compatibility condition with the product

$$(P^{\alpha\beta}{}_{\zeta\eta})^* P^{\eta\zeta}{}_{\gamma\delta} = P^{\beta\alpha}{}_{\gamma\delta}$$

is satisfied provided $Q^{\alpha\beta}{}_{\gamma\delta}$ is real.

To simplify the formulae we introduce the notation

$$[\lambda_\alpha, \lambda_\beta] = \Lambda_{\alpha\beta}. \tag{16}$$

We can then write (7) in the form

$$\Lambda_{\alpha\beta} + 2i\bar{k}\Lambda_{\gamma\delta}\, Q_-^{\gamma\delta}{}_{\alpha\beta} = K_{\alpha\beta} + \lambda_\gamma (F^\gamma{}_{\alpha\beta} - 2i\bar{k}\lambda_\delta Q_-^{\gamma\delta}{}_{\alpha\beta}). \tag{17}$$

This implies that to lowest order we can rewrite (7) as two independent equations

$$\Lambda_{\alpha\beta} = K_{-\alpha\beta} + \lambda_\gamma F^\gamma_{-\alpha\beta} - 2i\bar{k}\lambda_\gamma\lambda_\delta Q_-^{\gamma\delta}{}_{\alpha\beta}, \tag{18}$$

$$0 = K_{+\alpha\beta} + \lambda_\gamma F^\gamma_{+\alpha\beta} - 2i\bar{k}\lambda_\gamma\lambda_\delta Q_+^{\gamma\delta}{}_{\alpha\beta}. \tag{19}$$

This is the form which we shall use. If Equation (19) is non-trivial then one can substitute into the third term on the right-hand side the expression (18) for the commutator:

$$2i\bar{k}\lambda_\gamma\lambda_\delta Q^{\gamma\delta}{}_{\alpha\beta} = i\bar{k}\Lambda_{\gamma\delta} Q^{\gamma\delta}{}_{\alpha\beta}.$$

From the definition (10) it follows that

$$C^\gamma{}_{\alpha\beta} = F^\gamma{}_{\alpha\beta} - 4i\bar{k}\lambda_\delta Q_-^{\gamma\delta}{}_{\alpha\beta}. \tag{20}$$

We must choose the $\Lambda_{\alpha\beta}$ so that for arbitrary $f \in \mathcal{A}$ in the classical limit when $f \to \tilde{f}$

$$[\tilde{e}_\alpha, \tilde{e}_\beta] f = \tilde{C}^\gamma{}_{\alpha\beta} \tilde{e}_\gamma f.$$

From the general considerations of Section 2 and in particular Equation (17) we have

$$[\Lambda_{\alpha\beta}, f] = F^\gamma{}_{\alpha\beta}[\lambda_\gamma, f] - 2i\bar{k} Q^{\gamma\delta}{}_{\alpha\beta}[\lambda_\gamma\lambda_\delta, f]$$

which we rewrite as

$$[e_\alpha, e_\beta] f = \tfrac{1}{2} C^\gamma{}_{\alpha\beta} e_\gamma f + \tfrac{1}{2} e_\gamma f C^\gamma{}_{\alpha\beta}. \tag{21}$$

We shall assume that

$$F^\gamma{}_{\alpha\beta} = 0.$$

In the classical limit we have

$$C^\alpha{}_{\beta\gamma} \to \tilde\theta^\alpha_\mu \tilde e_{[\beta} \tilde e^\mu_{\gamma]}.$$

The gauge-fixing condition can be written to the classical approximation

$$e_\gamma(\theta^\gamma_\mu e_{[\alpha} e^\mu_{\beta]}) = 0.$$

We assume that the noncommutative expression is the same to within a reordering of the factors. To determine the order we consider the Jacobi identity

$$[[\lambda_\alpha, \lambda_\beta], x^\mu] + [[\lambda_\beta, x^\mu], \lambda_\alpha] + [[x^\mu, \lambda_\alpha], \lambda_\beta] = 0.$$

Using (21) we can rewrite this as

$$\tfrac{1}{2} C^\gamma{}_{\alpha\beta} e^\mu_\gamma + \tfrac{1}{2} e^\mu_\gamma C^\gamma{}_{\alpha\beta} = e_{[\alpha} e^\mu_{\beta]}.$$

We define then $C^\gamma{}_{\alpha\beta}$ to be a solution of this equation. Because of the standard ordering problems familiar from quantum mechanics the solution will not be unique.

To lowest order one can write

$$C^\gamma{}_{\alpha\beta} = \theta^\gamma_\mu e_{[\alpha} e^\mu_{\beta]} - \tfrac{1}{2}[\theta^\gamma_\mu, e_{[\alpha} e^\mu_{\beta]}] = \tfrac{1}{2}(\theta^\gamma_\mu e_{[\alpha} e^\mu_{\beta]} + e_{[\alpha} e^\mu_{\beta]} \theta^\gamma_\mu).$$

4 The connection

It is necessary [1] to introduce a flip operation

$$\sigma : \ \Omega^1(\mathcal{A}) \otimes \Omega^1(\mathcal{A}) \to \Omega^1(\mathcal{A}) \otimes \Omega^1(\mathcal{A})$$

to define the reality condition and the Leibniz rules. If we write

$$S^{\alpha\beta}{}_{\gamma\delta} = \delta^\beta_\gamma \delta^\alpha_\delta + i\!\!k T^{\alpha\beta}{}_{\gamma\delta}$$

we find that a choice [8] of connection which is torsion-free, and satisfies all Leibniz rules is given by

$$\omega^\alpha{}_\beta = \tfrac{1}{2} F^\alpha{}_{\gamma\beta} \theta^\gamma + i\!\!k \lambda_\gamma T^{\alpha\gamma}{}_{\delta\beta} \theta^\delta. \tag{22}$$

The relation

$$\pi \circ (1 + \sigma) = 0$$

must hold [8,9] to assure that the torsion be a bilinear map.

To all orders one has

$$\omega^\alpha{}_{\eta\zeta} P^{\eta\zeta}{}_{\beta\gamma} = \tfrac{1}{2} C^\alpha{}_{\beta\gamma}. \tag{23}$$

This is the usual relation between the Ricci-rotation coefficients and the Levi-Civita connection. Using it one can deduce (10) from (22). One can also write (22) in the form

$$2\omega^\alpha{}_{\delta\beta} = C^\alpha{}_{\delta\beta} + i\!\!k \lambda_\gamma T^{\alpha\gamma}{}_{(\delta\beta)} - i\!\!k \lambda_\gamma T^{\alpha\gamma}{}_{(\eta\zeta)} Q^{\eta\zeta}{}_{\delta\beta}$$

which includes the part symmetric in the second two indices and which must be determined by the condition that the connection be metric. In terms of the coefficients $P^{\alpha\beta}{}_{\gamma\delta}$ the relation can be written in the form

$$T^{\alpha\beta}{}_{\eta\zeta}P^{\eta\zeta}{}_{\gamma\delta} + Q^{(\alpha\beta)}{}_{\gamma\delta} = 0. \tag{24}$$

To lowest order this becomes

$$Q_-^{\alpha\beta}{}_{\gamma\delta} = -\tfrac{1}{4}T^{\alpha\beta}{}_{[\gamma\delta]}, \tag{25}$$

$$T^{\alpha\beta}{}_{\gamma\delta} = -2(Q_-^{\alpha\beta}{}_{\gamma\delta} - Q_{+(\gamma}{}^{\beta}{}_{\delta)}{}^{\alpha}). \tag{26}$$

The symmetric part of $T^{\alpha\beta}{}_{\gamma\delta}$ has been here fixed by the condition that the connection be metric.

Under a change of frame basis the coefficients of the spin connection also change. We mention only the linear approximation. If

$$\theta'^{\alpha} = \theta^{\alpha} - H^{\alpha}{}_{\beta}\theta^{\beta}$$

then

$$C'^{\alpha}{}_{\beta\gamma} = C^{\alpha}{}_{\beta\gamma} + D_{[\beta}H^{\alpha}{}_{\gamma]}.$$

The only restriction on $H^{\alpha}{}_{\beta}$, apart from the condition that it be small and antisymmetric, is that it must leave the condition (11) invariant or impose it if it is not satisfied. If we treat the λ_α as the components of a 1-form $-\theta$, which they are, and take their covariant derivative as if they formed an ordinary covector, which they do not, then we find that

$$D_{[\alpha}\lambda_{\beta]} - 2[\lambda_\alpha, \lambda_\beta] - C^{\gamma}{}_{\alpha\beta}\lambda_\gamma = 2K_{\alpha\beta}.$$

It is difficult to interpret this equation.

5 The metric

We shall suppose that $\mathcal{A}$ has a metric

$$g:\ \Omega^1(\mathcal{A}) \otimes \Omega^1(\mathcal{A}) \to \mathcal{A}. \tag{27}$$

In terms of the frame one can define the metric by the condition that

$$g(\theta^\alpha \otimes \theta^\beta) = g^{\alpha\beta}. \tag{28}$$

The $g^{\alpha\beta}$ are taken to form an arbitrary complex matrix which satisfies [23] the symmetry condition

$$P^{\alpha\beta}{}_{\gamma\delta}g^{\gamma\delta} = 0 \tag{29}$$

as well as the reality condition

$$g^{\beta\alpha} + i\hbar\!\!\!k T^{\alpha\beta}{}_{\gamma\delta}g^{\gamma\delta} = (g^{\beta\alpha})^*.$$

When constructing an algebra from a given classical geometry usually, but not necessarily, one starts with the matrix $g^{\alpha\beta}$ of standard Minkowski or euclidian

metric components. One must read the symmetry and reality conditions then as conditions on the maps π and σ.

If we write $g^{\alpha\beta} = \eta^{\alpha\beta} + i\bar{k} h^{\alpha\beta}$ then we find that to first order the symmetry and reality become respectively

$$h^{\alpha\beta} = -Q_{+}^{\alpha\beta}{}_{\gamma\delta}\eta^{\gamma\delta} = \frac{1}{2}T^{\alpha\beta}{}_{\gamma\delta}\eta^{\gamma\delta}.$$

If we require that the flip be an involution then we find to first order that

$$T^{\alpha\beta}{}_{\gamma\delta} + 2Q^{\alpha\beta}{}_{\gamma\delta} = 0.$$

The various reality conditions [22,23] imply also that

$$(Q^{\alpha\beta}{}_{\gamma\delta})^* = Q^{\alpha\beta}{}_{\gamma\delta} + o(i\bar{k}), \qquad (T^{\alpha\beta}{}_{\gamma\delta})^* = T^{\alpha\beta}{}_{\gamma\delta} + o(i\bar{k}).$$

The sum of two idempotents is also an idempotent if the two terms are orthogonal.

The condition that the product be a projector implies that it be hermitian with respect to the usual inner product on the tensor product:

$$g^{\alpha\gamma}g^{\beta\delta} = (\theta^\alpha \otimes \theta^\beta, \theta^\gamma \otimes \theta^\delta).$$

Therefore we have the condition

$$P^{\alpha\beta}{}_{\gamma\delta}g^{\gamma\zeta}g^{\delta\eta} = P^{\zeta\eta}{}_{\gamma\delta}g^{\gamma\alpha}g^{\delta\beta} \tag{30}$$

for $P^{\alpha\beta}{}_{\gamma\delta}$ and the metric. A weaker condition is the orthogonality condition

$$P^{\alpha\beta}{}_{\gamma\delta}g^{\gamma\zeta}g^{\delta\eta}(\delta^\mu_\zeta\delta^\nu_\eta + S^{\mu\nu}{}_{\zeta\eta}) = 0.$$

With the Ansatz we shall use this second condition is an identity.

The connection is compatible with the metric if

$$T^{\alpha\gamma}{}_{\delta\epsilon}g^{\epsilon\beta} + T^{\beta\gamma}{}_{\delta\epsilon}g^{\alpha\epsilon} + i\bar{k}T^{\beta\gamma}{}_{\epsilon\zeta}g^{\eta\zeta}T^{\alpha\epsilon}{}_{\delta\eta} = 0. \tag{31}$$

To first order this simplifies to the usual condition

$$T^{(\alpha\gamma}{}_{\delta}{}^{\beta)} = o(i\bar{k}). \tag{32}$$

The index was lifted here with the lowest-order, symmetric part of the metric.

6 Speculations

It is tempting to suppose that to lowest order at least, in a semi-classical approximation, there is an analogue of Darboux's lemma and that it is always possible to choose generators which satisfy commutation relations of the form (2) with the right-hand in the center. However the example we shall examine in detail shows that this is not always the case. Having fixed the generators, the manifestations of curvature would be found then in the form of the frame. The two sets of generators x^μ and λ_α satisfy, under the assumptions we make, three sets of equations.

The commutation relations (2) for the position generators x^μ and the associated Jacobi identities permit one definition of the algebra. The commutation relations for the momentum generators permit a second definition. The conjugacy relations assure that the two descriptions concern the same algebra. We shall analyze these identities later using the example to show that they have interesting non-trivial solutions.

The problem of gauge invariance and the algebra of observables is a touchy one upon which we shall not dwell. It is obvious that not all of the elements of $\mathcal{A}$ are gauge invariant but not that all observables are gauge-invariant. One of the principles of the theory of general relativity is that all (regular) coordinates systems or frames are equal. In the noncommutative case one finds that some are more equal than others. If one quantize a space-time using two different moving frames one will obtain two different differential calculi, although the two underlying algebras might be the same. This is equivalent to the fact that the canonical transformations of a commutative phase space are a very special set of phase-space coordinate transformations. It can also be expressed as the fact that the Poisson structure which remains on space-time as the commutative limit of the commutation relations breaks Lorenz invariance. In the special case where the H^α_β are constants then the two quantized frames will be also equivalent. Since we have decided to work only with algebras whose centers are trivial the converse will also be true. For a discussion of Poisson structures on curved manifolds we refer to Fedosov [24]. Since we are interested in finding the 'simplest' differential calculus, one of the aspects of the problem is the choice of 'correct' moving frame to start with.

One possible method of looking for a solution is to consider a manifold V embedded in $\mathbb{R}^d$ for some d with the commutation relations

$$[y^i, y^j] = i\!\!k J^{ij}, \qquad J^{ij} \in \mathbb{R}.$$

This will induce a symplectic structure on V which is intimately related to the one we shall exhibit in the following sections. The details of this have yet to be investigated. Let the larger algebra be $\mathcal{B}$. It has a natural differential calculus defined by imposing the condition $[y^i, dy^j] = 0$ that the differentials of the generators be a frame. It follows that the associated metric is flat. The projection

$$\Omega^*(\mathcal{B}) \longrightarrow \Omega^*(\mathcal{A})$$

would yield a solution to the problem but it is not necessarily easier to find. In fact a similar situation arises in one of the possible definitions of a differential calculus as a quotient of the universal differential calculus by a differential ideal. In that case the projection is strictly equivalent to the calculus. One could also consider the problem of finding the metric as an evolution equation in field theory in the sense that one can pass from the Schrödinger picture to the Heisenberg picture with the help of an evolution hamiltonian.

It is interesting to notice how the old Kaluza-Klein idea of gauge transformations as coordinate transformations appears here. Gauge transformations are inner automorphisms of the algebra with respect to some unitary (pseudo-)group

$\mathcal{G}_G \subset \mathcal{A}$ of elements; the complete dynamical evolution of the system can be described as an involution with respect to one unitary element $U = e^{iHt}$ of a (pseudo-)group $\mathcal{G}_H \subset \mathcal{A}$ of elements of $\mathcal{A}$, just as in quantum field theory. The difference lies in the 'size' of the subalgebra $\mathcal{A}_G$ in which $\mathcal{G}$ takes its values, as can be measured for example by the dimension of the commutant of the subalgebra generated by it; whereas in general $\dim(\mathcal{A}'_G) = \dim(\mathcal{A}')$, since gauge transformations are relatively unimportant, in general $\dim(\mathcal{A}'_H) = 0$. A topological field theory has $\dim(\mathcal{A}'_G) = \dim(\mathcal{A}'_H)$.

A Riemann-flat solution to the problem is given by choosing

$$e^\mu_\alpha = \delta^\mu_\alpha, \qquad K_{\alpha\beta} = -\frac{1}{i\kern-.3em k} J^{-1}_{\alpha\beta} \in \mathcal{Z}(\mathcal{A}).$$

We have introduced the inverse matrix $J^{-1}_{\alpha\beta}$ of $J^{\alpha\beta}$; we must suppose the Poisson structure to be non-degenerate: $\det J^{\alpha\beta} \neq 0$. The relations can be written in the form

$$\lambda_\alpha = -K_{\alpha\mu} x^\mu, \qquad [\lambda_\alpha, \lambda_\beta] = K_{\alpha\beta}. \tag{33}$$

This structure is flat according to our definitions.

The most natural Ansatz for the coefficient array Q would seem to be of the form

$$Q_-^{\alpha\beta}{}_{\gamma\delta} = \tfrac{1}{4} k^{(\alpha} Q^{\beta)}_{[\gamma} k_{\delta]}$$

with k_α a principle null vector and Q^β_γ a matrix. Because of the symmetry and reality conditions on the metric one must suppose that

$$g_{\alpha\beta} Q_-^{\alpha\beta}{}_{\gamma\delta} = 0.$$

This equation will be satisfied if k^α is an eigenvector of Q^β_α. We set

$$Q^\beta_\alpha k^\alpha = q k^\beta$$

and we conclude that

$$k_\alpha Q_-^{\alpha\beta}{}_{\gamma\delta} = 0.$$

From (20) we obtain the expression

$$C^\alpha{}_{\beta\gamma} \lambda_\alpha = F^\alpha{}_{\beta\gamma} \lambda_\alpha - i\kern-.3em k \mu^2 \lambda_\alpha \lambda_\delta k^{(\alpha} Q^{\delta)}_{[\beta} k_{\gamma]}$$

This leads to an expression for τ which is a constant multiple of $\lambda_\alpha k^\alpha$:

$$\tau = -2 i\kern-.3em k \mu^2 \lambda_\alpha k^\alpha.$$

We would like τ to play the role of time and so must impose the condition that $e_a \tau = 0$ and $e_0 \tau \neq 0$. One sees, with $F^\alpha{}_{\beta\gamma} = 0$ and to lowest order, that

$$e_\alpha \tau \propto k^\beta [\lambda_\alpha, \lambda_\beta] = k^\beta K_{\alpha\beta}.$$

The null vector must be chosen to have the correct relation also with the symplectic form. It seems difficult however to work with this Ansatz and so we shall opt for a simpler if less elegant one.

We shall find it convenient to consider a curved geometry as a perturbation of a noncommutative flat geometry. The measure of noncommutativity is the parameter $\bar{k}$; the measure of curvature is the quantity μ^2. There are two special interesting limits. If we keep $\bar{k}\mu^2$ small but fixed then we can let $\bar{k} \to 0$ or $\bar{k} \to \infty$. The former (latter) corresponds to a 'small' ('large') universe filled with 'small' ('large') cells. The number of cells is given by $(\bar{k}\mu^2)^{-1}$. We can assume the flat-space limit to have commutation relations of the form (1) with

$$J^{\mu\nu} = J^{\mu\nu}_{(0)}(1 + o(i\bar{k}\mu^2)).$$

Finally we mention that according to the standard definition of curvature, even corrected to account for the bimodule structure of the module of 1-forms, all the geometries we consider here have constant curvature [23]. One of the motivations for considering the examples is the hope that they will lead to a definition of curvature which is both bilinear as a map and also in some sense time-dependent.

7 The Kasner geometry

The Kasner metric is of Petrov type I and has four distinct principal null vectors. The limiting Poisson structure defines an additional two principle null vectors. We must also choose the frame so that it is in some way adapted to these vectors. A major problem is to possess a criterium by which one can decide if the frame is well-chosen. One obvious condition is that the frame components of all principal null vectors lie in the center of the algebra.

Choose a symmetric matrix $P = (P^a_b)$ of real numbers. A moving frame for the Kasner metric is given by

$$\tilde{\theta}^0 = d\tilde{t}, \qquad \tilde{\theta}^a = d\tilde{x}^a - P^a_b \tilde{x}^b \tilde{t}^{-1} d\tilde{t}. \tag{34}$$

The 1-forms $\tilde{\theta}^\alpha$ are dual to the derivations

$$\tilde{e}_0 = \tilde{\partial}_0 + P^i_j \tilde{x}^j \tilde{t}^{-1} \tilde{\partial}_i, \qquad \tilde{e}_a = \tilde{\partial}_a$$

of the algebra $\mathcal{A}$. The space $\mathcal{X}$ of all derivations is free of rank 4 as an $\mathcal{A}$-module and the $\tilde{e}_\alpha$ form a basis. The Lie-algebra structure of $\mathcal{X}$ is given by the commutation relations

$$[\tilde{e}_a, \tilde{e}_0] = \tilde{C}^b{}_{a0} \tilde{e}_b, \qquad [\tilde{e}_a, \tilde{e}_b] = 0 \tag{35}$$

with

$$\tilde{C}^b{}_{a0} = P^b_a \tilde{t}^{-1}.$$

For fixed time it is a solvable Lie algebra which is not nilpotent. We have written the frame in coordinates which are adapted to the asymptotic condition. There is a second set which will be convenient with space coordinates x'^a given in matrix notation by

$$x'^a = (t^{-P} x)^a.$$

The frame can be then written, again in matrix notation, with space components in the form

$$\theta^a = (t^P d(t^{-P} x))^a = (t^P dx')^a.$$

This transformation can be considered as an inner automorphism of the differential

$$d \mapsto d' = t^P \circ d \circ t^{-P}$$

and it would transform the differential calculus into a new one with a flat metric,

The expression for $\tilde{C}^b{}_{a0}$ contains no parameters with dimension but it has the correct physical dimensions. Let G_N be Newton's constant and μ a mass such that $G_N\mu$ is a length scale of cosmological order of magnitude. As a first guess we would like to identify the length scale determined by $\bar{k}$ with the Planck scale: $\hbar G_N \sim \bar{k}$ and so we have $\bar{k} \sim 10^{-87}\mathrm{sec}^2$ and since μ^{-1} is the age of the universe we have $\mu \sim 10^{-17}\mathrm{sec}^{-1}$. The dimensionless quantity $\bar{k}\mu^2$ is given by $\bar{k}\mu^2 \sim 10^{-120}$. In the Kasner case the role of μ is played by $\tilde{t}^{-1}$ at a given epoch $\tilde{t}_0$.

We saw, and we shall see below, that the spectrum of the commutator of two momenta is the sum of a constant term of order $\bar{k}^{-1}$ and a 'gravitational' term of order $\mu\tilde{t}^{-1} = \bar{k}^{-1} \times (\bar{k}\mu)\tilde{t}^{-1}$. So the gravitational term in the units we are using is relatively important for $\tilde{t} \lesssim \bar{k}\mu$. The existence of the constant term implies that the gravitational field is not to be identified with the noncommutativity *per se* but rather with its variation in space and time.

The components of the curvature form are given by

$$\tilde{\Omega}^a{}_0 = (P^2 - P)^a_b \tilde{t}^{-2}\tilde{\theta}^0\tilde{\theta}^b, \tag{36}$$

$$\tilde{\Omega}^a{}_b = -\tfrac{1}{2}P^a_{[c}P_{d]b}\tilde{t}^{-2}\tilde{\theta}^c\tilde{\theta}^d. \tag{37}$$

The curvature form is invariant under a uniform scaling of all coordinates. The Riemann tensor has components

$$\tilde{R}^a{}_{0c0} = (P^2 - P)^a_b \tilde{t}^{-2}, \qquad \tilde{R}^a{}_{bcd} = P^a_{[c}P_{d]b}\tilde{t}^{-2}.$$

The vacuum field equations reduce to the equations

$$\mathrm{Tr}(P) = 1, \qquad \mathrm{Tr}(P^2) = 1.$$

If p_a are the eigenvalues of the matrix P^a_b there is a 1-parameter family of solutions given by

$$p_a = \frac{1}{1+\omega+\omega^2}\,(1+\omega,\ \omega(1+\omega),\ -\omega)\,. \tag{38}$$

The most interesting value is $\omega = 1$ in which case

$$p_a = \tfrac{1}{3}(2,\ 2,\ -1).$$

The curvature invariants are proportional to $\tilde{t}^{-2}$; they are singular at $\tilde{t} = 0$ and vanish as $\tilde{t} \to \infty$.

The values $p_a = c$ for the three parameters are also of interest. The Einstein tensor is given by

$$\tilde{G}^0_0 = -3c^2\tilde{t}^{-2}, \qquad \tilde{G}^a_b = -c(3c-2)\delta^a_b\tilde{t}^{-2}.$$

For the value $c = 2/3$ the space is a flat FRW with a dust source given by

$$\tilde{T}_{00} = -\frac{1}{8\pi G_N}\tilde{G}_{00} = \frac{1}{6\pi G_N}.$$

For $c = 1/3$ the space is Einstein with a time-dependent cosmological 'constant' which we interpret as due to the presence of the supplementary noncommutative dimensions. We start then with a Kasner solution in dimension 4 and we add a fuzzy structure which forces us out of the 1-parameter family of vacuum solutions into a family of solutions with similar properties except for the existence of a non-zero value for Λ_c. We interpret the time dependence of the 'constant' as due to the time variation of the internal structure which gave rise to it but we cannot write explicitly a formulae as one does in the case of an finite-dimensional manifold as internal structure or in the case [20] of a finite-dimensional noncommutative structure. We have in the present case an infinite-dimensional noncommutative algebra and we would expect rather Λ_c to yield information about the algebra than the inverse.

To illustrate the notation one can analyze the non-commutative version of the region in parameter space around the flat solution given by $\omega = -1$. The local Lorentz rotation which makes this explicit is equivalent to a change of coordinates. If we choose the $\tilde{z}$-axis along the vector with the non-vanishing eigenvalue then the transformation is given by

$$\tilde{t}' = \tilde{z}\sinh(\tilde{t}/\tilde{z}), \qquad \tilde{z}' = \tilde{z}\cosh(\tilde{t}/\tilde{z}).$$

It follows that $\tilde{t}^2 = \tilde{t}'^2 - \tilde{z}'^2$; the origin of the Kasner time coordinate, exactly at the flat-space values of the parameters and because of the singular nature of the transformation, becomes a null surface.

8 The momentum generators

There are four sets of Jacobi identities to satisfy, depending on how many momentum factors are present. We shall start with those depending only on the momenta since they were analyzed in detail in the general discussion. If we assume that $F^{\alpha}{}_{\beta\gamma} = 0$ and that the noncommutative extension has the same form as the commutative limit then from the results of Section 2, in the Kasner case, we find that the commutator $\Lambda_{\alpha\beta}$ is of the form

$$\Lambda_{ab} = K_{ab}, \tag{39}$$

$$\Lambda_{0a} = K_{0a} + \tfrac{1}{2}C^{b}{}_{0a}\lambda_b. \tag{40}$$

The second term on the right-hand side of the second equation can also be written as

$$\tfrac{1}{2}C^{b}{}_{0a}\lambda_b = -2i\bar{k}\lambda_b\lambda_c Q^{bc}{}_{0a}.$$

We shall suppose that for some matrix P^a_b

$$P^{bc}{}_{a0} = \tfrac{1}{4}\mu^2 P^{(b}_a P^{c)}_d k'^d$$

so that

$$C^b{}_{a0}\lambda_b = -i\bar{k}\mu^2\lambda_{(b}\lambda_{c)}P^b_a P^c_d k'^d.$$

From the classical limit we can infer therefore that for some τ

$$P^b_a(\tau\lambda_b + \lambda_b\tau) = -2i\bar{k}\lambda_{(b}\lambda_{c)}P^b_a P^c_d k'^d.$$

We have argued in fact that in general the commutative limit should be flat and so this relation should be expected to be satisfied only in the limiting case $\omega \to 1$. We define

$$\tau = -2i\bar{k}\mu^2 k'^a P^b_a \lambda_b \tag{41}$$

to obtain a consistent relation.

We recall that the right-hand side of (39) as well as the first term on the right-hand side of (402) diverge as $\bar{k} \to 0$. To stress this fact we write

$$K_{0a} = (i\bar{k})^{-1} l_a, \qquad K_{ab} = (i\bar{k})^{-1}\epsilon_{abc}k^c$$

with two space-like vectors l_a and k^a. These constitute a sort of 'vacuum-energy density'; they imply that in the flat limit the commutation relations are not necessarily trivial. We shall require that τ depend only on time, that

$$e_a\tau = 0.$$

From the definition it follows that

$$e_a\tau = -2i\bar{k}\mu^2 k'^c P^b_c K_{ab}$$

and therefore that

$$k'^c P^b_c K_{ab} = 0.$$

We choose accordingly

$$k^a = P^a_b k'^b, \qquad \tau = -2i\bar{k}\mu^2 k^a\lambda_a.$$

It will be convenient to introduce the matrix

$$W(t) = \exp -\tfrac{1}{2}P\int_1^t \tau dt$$

solution to the differential equation

$$\dot{W} + \tfrac{1}{2}P\tau W = 0.$$

The three eigenvalues of this matrix are the integrating factors for the frame. Once we have found a particular solution λ_a to Equation (40) then the most general solution is obtained by adding to it a term of the form $W^b_a\mu_b$, with μ_b a triplet of operators which commute with t.

A natural connection between the symplectic and metric structures is that the vector k^a dual to K_{ab} be an eigenvector of P^a_b. Let p be the corresponding eigenvalue. It follows that if we multiply both sides of (40) by $-2i\mu^2k^a$ we find then for τ the equation

$$\dot{\tau} + \tfrac{1}{2}p\tau^2 + 2\mu^2 l_a k^a = 0. \tag{42}$$

This equation transforms under the action of the 'duality' transformation

$$\tau \mapsto -\mu^2\tau^{-1}$$

into

$$\dot{\tau} + 2l_a k^a \tau^2 + \tfrac{1}{2}p\mu^2 = 0. \tag{43}$$

If one neglect the higher-order terms the behaviour of these solutions near the singularity is respectively

$$\tau \simeq \frac{2}{pt}, \qquad \tau \simeq \frac{1}{2l_a k^a t}.$$

In the commutative limit $W \to \tilde{W}$ with

$$\tilde{W} = \tilde{t}^{-P/p}.$$

It would appear here that the classical limit is not a smooth one. This point is not clear. We cannot expect the two solutions to behave similarly near the singularity; indeed we hope to eventually find a 'smooth' noncommutative analog. Therefore $\tau - t^{-1}$ should diverge near the singularity. On the other hand in the asymptotic region the noncommutative extension has a cosmological constant and is thus not comparable with the classical solution either. The best we can claim here according to Equation (44) (with the plus sign) is that near the flat solution with $p \simeq 1$ and in the intermediate region with $\tau \simeq \mu$ the noncommutative solution behaves like the classical analog.

It will be convenient to impose the condition

$$4l_a k^a = \pm p$$

and write Equation (42) as

$$\dot{\tau} + \tfrac{1}{2}p(\tau^2 \pm \mu^2) = 0. \tag{44}$$

Equation (44) is a degenerate form of the Riccati equation and can be completely solved. With the plus sign the function

$$\tau = \mu \cot(p\mu t/2) \tag{45}$$

is a solution. With this solution the expression for W becomes

$$W = \sin^{-P/p}(p\mu t/2).$$

Using the Jacobi identity as a Leibniz rule, we find the differential equation

$$\dot{\Lambda}_{ab} = -\tfrac{1}{2}P^c_{[a}\Lambda_{cb]} \tag{46}$$

for the commutators Λ_{ab}. This equation can solved to yield the expression

$$\Lambda_{ab} = W^c_a W^d_b K_{cd} = (\sin^{-P/p})^c_a (\sin^{-P/p})^d_b K_{cd}.$$

In the preferred coordinate system, with the z-axis along the direction of the distinguished eigenvector k^a, the only nonvanishing element is

$$\Lambda_{12} = \sin^{(p-1)/p}(p\mu t/2) K_{12}.$$

We shall suppose that $g^{ab} l_b$ and k^a are parallel. We can solve then the system of equations (40). We find that for $a = 1, 2$

$$\lambda_a = W_a^b \mu_b \tag{47}$$

with μ_b a doublet of operators which commute with t. The third element λ_3 is a constant multiple of τ.

To clarify this a bit one can consider a perturbative solution near flat space, that is with the parameter ω approximately equal to the flat-space value $\omega = -1$. Since we have claimed that non-trivial commutation relations and a curved metric are in fact two aspects of the same reality, we allow ourselves the freedom to identify the difference $\omega + 1$ with the parameter $\bar{k}\mu^2$. We write then $\omega = -1 + \bar{k}\mu^2$ and expand P in a power series:

$$P = P_0 + \bar{k}\mu^2 P_1 + (\bar{k}\mu^2)^2 P_2 + \cdots .$$

The coefficients can be written in the form

$$P_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

To first order the expansion of the right-hand side of (40) is given by

$$[\lambda_0, \lambda_1] = -\tfrac{1}{2}\bar{k}\mu^2 \tau \lambda_1, \tag{48}$$

$$[\lambda_0, \lambda_2] = \tfrac{1}{2}\bar{k}\mu^2 \tau \lambda_2, \tag{49}$$

$$[\lambda_0, \lambda_3] = \frac{1}{i\bar{k}} l_3 - \tfrac{1}{2}\tau \lambda_3. \tag{50}$$

The structure of the classical algebra of derivations is given by (35); the noncommutative generalization is an algebra defined by the relations (39) and (40) with however a different right-hand side for the former.

It is possible to represent $\mathcal{A}$ as a tensor product of two Heisenberg algebras. We introduce μ_α with

$$[\mu_\alpha, \mu_\beta] = K_{\alpha\beta}$$

and we define $\mathcal{A}_{12}$ ($\mathcal{A}_{30}$) to be the algebra generated by (μ_1, μ_2) ((μ_3, μ_0)). We define an embedding

$$0 \to \mathcal{A}_{30} \xrightarrow{\phi} \mathcal{A}$$

by setting

$$\lambda_0 = \phi(\mu_0) = \mu_0, \qquad \lambda_3 = \phi(\mu_3) = -(2i\bar{k}\mu)^{-1}\cot(\tfrac{1}{2}p\mu\mu_3).$$

The image of $\mathcal{A}_{30}$ is an 2-sided ideal of $\mathcal{A}$ and the projection

$$\mathcal{A}_{12} \otimes \mathcal{A}_{30} \stackrel{\phi}{\to} \mathcal{A}_{12} \to 0$$

is defined by setting, for $a = 1, 2$

$$\lambda_a = \phi_a(\mu_a, \mu_3) = W_a^b \mu_b. \tag{51}$$

We have represented $\mathcal{A}$ as the direct product of $\mathcal{A}_{12}$ and $\mathcal{A}_{30}$. The 'canonical transformation' $\mu_\alpha \mapsto \lambda_\alpha$ is partially given by a change in momenta defined by a transendental function; it is highly nonlocal. There is evidence of what could be considered a noncommutative Darboux theorem. In the formal analogy with quantum field theory the μ_α could be perhaps considered 'bare' momenta and the λ_α the corresponding 'dressed' momenta. There has however been no use made of the latter and we shall wait until we have introduced the position generators and can speak of 'bare' and 'dressed' fields to pursue this analogy.

9 The position generators

The conjugacy relations which define the Kasner metric are most easily defined using the second set x'^a of coordinates introduced in Section 7. They are given by

$$\begin{aligned} &[\lambda_0, t'] = 1,\ [\lambda_0, x'^j] = 0, \\ &[\lambda_a, t'] = 0,\ [\lambda_a, x'^j] = (W^{-2})^j_a. \end{aligned} \tag{52}$$

This set of equations constitutes a relation between the coordinate generators and the momenta generators. Using them one can find differential equations for the $J^{\mu\nu}$:

$$\begin{aligned} &[\lambda_0, J'^{ij}] = 0,\ [\lambda_a, J'^{ij}] = (PW^{-2})^{[i}_a \tau J'^{0j]}, \\ &[\lambda_0, J'^{0j}] = 0,\ [\lambda_a, J'^{0j}] = 0. \end{aligned} \tag{53}$$

These can be solved immediately to yield the expresssions

$$J'^{ij} = S'^{ij}_{(0)} + L'^{ij}, \qquad L'^{ij} = P^{[i}_k x'^k \tau J'^{0j]}, \qquad J'^{0i} = J'^{0j}_{(0)}, \tag{54}$$

for the commutator as a sum of a constant 'spin' S'^{ij} and 'orbital momentum' L'^{ij}. They can also be written in terms of the original coordinates x^i, without due regard to hermiticity, as

$$J^{0i} = [t, (W^2 x')^i] = (W^2 J')^{0i} = (W^2)^i_j J'^{0j}_{(0)}.$$

and a rather more complicated expression

$$J^{ij} = [(W^2 x')^i, (W^2 x')^j]$$

for the angular momentum. The components of L'^{ij} behave as t^{-2}

If we choose the z-axis so that $\bar{k} J'^{0i} = (0, 0, ih)$ and neglect the 'spin' then we find the Lie algebra structure

$$[x', y'] = 0, \qquad [y', z'] = ihy\prime, \qquad [z', x'] = -ihx'$$

of the solvable subgroup of $SL(2,\mathbb{R})$. This algebra appears also in a noncommutative version of the Lobachevski plane. We refer elsewhere [20] for a discussion of this as well as for the reference to the original literature.

It is difficult to appreciate the significance of these components since at the origin the coordinate system becomes singular. The frame components

$$I^{\alpha\beta} = \theta^\alpha_\mu \theta^\beta_\nu J^{\mu\nu}$$

might be considered more significant. One finds that

$$I^{0a} = (W^2 J_{(0)})^{0a}, \qquad I^{ab} = S^{ab} = (W^4 S_{(0)})^{ab}$$

provided one uses the 'frame components' of the coordinate (generator)

$$x^a = (W^2)^a_i x'^i = x^i.$$

We can choose as only non-vanishing components $S^{12}_{(0)}$ and J^{03} in which case near the singularity we find that

$$I^{a3} \sim t^{-2}.$$

The 'spin' has to lowest order the same time dependence as the curvature. the singularity is infinitely fuzzy. In this sense we have resolved the point singularity.

We have noticed that when $p_a = 2/3(1,1,1)$ the space becomes the flat FRW solution with a pressure-free dust as source. The matrix W is a multiple of the unit matrix and we find for the momenta commutators

$$\Lambda_{ab} \sim t^{-2}$$

The standard coordinates one uses are the analogs of the x'^i coordinates introduced in Section 7 for the Kasner metric. In this case the position generators have commutators which near the origin behave as

$$L'^{ij} = \tfrac{2}{3} x'^{[i} t^{-1} J'^{0j]}.$$

The covariant derivatives $D_\gamma I^{\alpha\beta}$ of the 'spin tensor' are given as

$$D_0 I^{0a} = e_0 I^{0a} = P^a_b \tau J^{b0}, \quad D_b I^{0a} = 0,$$
$$D_0 I^{ab} = e_0 I^{ab} = \tau P^{[a}_c I^{bc]}, \; D_c I^{ab} = e_c I^{ab} + \omega^{[b}{}_{a0} J^{0c]} = \tau P^{[a}_c I^{0b]}.$$

We find then a 'Maxwell field strength' $F_{\alpha\beta} = \mu^2 I_{\alpha\beta}$ obeying 'Maxwell's equations' with a current j_E defined by

$$j^0_E = 0, \qquad j^a_E = D_\alpha I^{\alpha a} = \tau\mu^2 I^{0a}.$$

There is also a 'magnetic monopole' density given by

$$j^a_M = \tfrac{1}{2} P^{[a}_b \tau I^{0b]} = \tfrac{1}{2} P^a_b j^b_E - j^a_E.$$

This relation recalls somewhat the relation between an magnetic field and an electric field in a uniformly moving Lorentz frame. If $B_a = \epsilon_{abc} v^b E^c$ depends only on time then

$$j_{Ma} = \dot{B}_a = \epsilon_{abc} v^b \dot{E}^c = \epsilon_{abc} v^b j^c_E.$$

One obtains the former from the latter by the replacement

$$\epsilon_{abc}\nu^b \mapsto \frac{1}{2}P_{ab} - g_{ab}$$

of an antisymmetric tensor by a symmetric one.

It is interesting to note than one can consider Equation (46) as the equation

$$\nabla_\alpha \theta^\beta_\nu = 0$$

which permits one to pass from a coordinate frame to an orthonormal frame. The covariant derivative here is with respect to the complete set of indices. This equation yields the relation between the Christophel symbols and the Ricci rotation coefficients.

Discusssion

In a subsequent publication we shall discuss the differential calculus. There is a well-defined if perhaps complicated algorithm to construct the frame starting with the momenta λ_α. Having 'blurred' the Kasner metric and deformed the resulting algebra we can now take the 'sharp' limit and see what we obtain. With the form of P^a_b we have the metric cannot be Ricci-flat but has an induced cosmological constant due to the noncommutativity [17]. The theory we are investigating has certain similarities with theories of the type called Kaluza-Klein. That is, the additional noncommutative structure can perhaps at least to a certain extent be assimilated to an effective commutative theory in higher dimensions. This means that even if one could define a curvature tensor in a satisfactory manner there is no reason to expect the Ricci tensor to vanish. We shall assume that to the lowest approximation the Ricci tensor of the total structure does vanish and we shall use the Ricci tensor of the four dimensions to elucidate the structure of the hidden dimensions.

Acknowledgments

One of the authors (JM) would like to thank M. Burić, S. Cho, G Fiore, E. Floratos, H. Grosse, J. Mourad, A. Sykora, S. Theisen, P. Tod and G. Zoupanos for interesting conversations. Preliminary work was to a large extent carried out while he was visiting the MPI in München. He is grateful to Julius Wess for financial support during this period. The results included here were presented by him at the II Summer School in Modern Mathematical Physics, Kopaonik, the XXI Workshop on Geometric Methods in Physics, Bialowieza, the Erwin Schrödinger Institute workshop on Noncommutative Geometry, Feynman Diagrams and Quantum Field Theory and the Euresco meeting on physics beyond the standard model in Portoroz, Slovenia.

References

1. M. Dubois-Violette, J. Madore, T. Masson, and J. Mourad, "On curvature in noncommutative geometry," *J. Math. Phys.* **37** (1996), no. 8, 4089–4102, q-alg/9512004.
2. M. Dubois-Violette, J. Madore, T. Masson, and J. Mourad, "Linear connections on the quantum plane," *Lett. Math. Phys.* **35** (1995), no. 4, 351–358, hep-th/9410199.
3. S. Cho, J. Madore, and K. S. Park, "Noncommutative geometry of the h-deformed quantum plane," *J. Phys. A: Math. Gen.* **31** (1998), no. 11, 2639–2654, q-alg/9709007.
4. G. Fiore and J. Madore, "The geometry of quantum euclidean spaces," *J. Geom. Phys.* **33** (2000) 257–287, math/9904027.
5. J. Madore, "On the resolution of space-time singularities," in Castellani *et. al.* [6], pp. 2419–2425. Torino, July 1999.
6. L. Castellani, F. Lizzi, and G. Landi, eds., "Noncommutative geometry and Hopf algebras in field theory and particle physics," *Euroconference on Noncommutative Geometry*. Torino, July 1999. *Int. J. Mod. Phys. B* **14** no. 22-23, World Scientific Publishing, 2000.
7. S. S. Abhyankar, *Resolution of Singularities of Embedded Algebraic Surfaces*. Springer Monographs in Mathematics. Springer-Verlag, second ed., 1998.
8. A. Dimakis and J. Madore, "Differential calculi and linear connections," *J. Math. Phys.* **37** (1996), no. 9, 4647–4661, q-alg/9601023.
9. J. Madore and J. Mourad, "Quantum space-time and classical gravity," *J. Math. Phys.* **39** (1998), no. 1, 423–442.
10. M. Heller and W. Sasin, "Non-commutative structure of singularities in general relativity," *J. Math. Phys.* **37** (1996) 5665.
11. E. Hawkins, "Noncommutative rigidity," http://arXiv.org/abs/math.qa/0211203.
12. F. Lizzi, G. Mangano, G. Miele, and M. Peloso, "Cosmological perturbations and short distance physics from noncommutative geometry," *J. High Energy Phys.* **06** (2002) 049, hep-th/0203099.
13. M. Gaul and C. Rovelli, "Loop quantum gravity and the meaning of diffeomorphism invariance," in *Towards Quantum Gravity*, J. Kowalski-Glikman, ed., no. 541 in Lect. Notes in Phys., pp. 277–324. Springer-Verlag, 1999. gr-qc/9910079. Polanica, February 1999.
14. P. Aschieri and L. Castellani, "An introduction to noncommutative differential geometry on quantum groups," *Int. J. Mod. Phys. A* **8** (1993) 1667–1706, hep-th/9207084.
15. A. H. Chamseddine, "Invariant actions for noncommutative gravity," hep-th/0202137.
16. M. Maceda, J. Madore, and D. C. Robinson, "Fuzzy-pp waves," *(to appear)* (2002) http://arXiv.org/abs/hep-th/0207225.
17. M. Maceda, J. Madore, P. Manousselis, and G. Zoupanos, "Can noncommutativity resolve the big-bang singularity?," hep-th/0306136.
18. J. Madore, "Kaluza-Klein aspects of noncommutative geometry," in *Differential Geometric Methods in Theoretical Physics*, A. I. Solomon, ed., pp. 243–252. World Scientific Publishing, 1989. Chester, August 1988.
19. B. L. Cerchiai, G. Fiore, and J. Madore, "Frame formalism for the N-dimensional quantum euclidean spaces," in Castellani *et. al.* [6], pp. 2305–2314. math.QA/0007044. Torino, July 1999.
20. J. Madore, *An Introduction to Noncommutative Differential Geometry and its Physical Applications*. No. 257 in London Mathematical Society Lecture Note Series. Cambridge University Press, second ed., 2000. 2nd revised printing.

21. J. Madore and L. A. Saeger, "Topology at the Planck length," *Class. and Quant. Grav.* **15** (1998) 811–826, `gr-qc/9708053`.
22. A. Connes, "Noncommutative geometry and reality," *J. Math. Phys.* **36** (1995) 6194.
23. G. Fiore and J. Madore, "Leibniz rules and reality conditions," *Euro. Phys. Jour. C* **17** (1998), no. 2, 359–366, `math/9806071`.
24. B. Fedosov, *Deformation Quantization and Index Theory*. Akademie-Verlag, Berlin, 1996.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 235)*
Portorož, Slovenia, July 12-17, 2003

The Multiple Point Principle: Realized Vacuum in Nature is Maximally Degenerate

Donald L. Bennett⋆[1] and Holger Bech Nielsen⋆⋆[2]

[1] Brookes Institute for Advanced Studies, Bøgevej 6, 2900 Hellerup, Denmark
[2] The Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

Abstract. We put forward the multiple point principle as a fine-tuning mechanism that can explain why some of the parameters of the standard model have the values observed experimentally. The principle states that the parameter values realized in Nature coincide with the surface (e.g. the point) in the action parameter space that lies in the boundary that separates the maximum number of regulator-induced phases (e.g., the lattice artifact phases of a lattice gauge theory). We argue that a mild form of non-locality - namely that embodied in allowing diffeomorphism invariant contributions to the action - seems to be needed for some fine-tuning problems. We demonstrate that the multiple point principle solution to fine-tuning has the very special property of avoiding the paradoxes that can arise in the presence of non-locality. The non-renormalizability of gravity suggests — in a manner reminiscent of baby universe theory — the presence of non-local effects without which the phenomenologically observed high degree of flatness of spacetime would seem mysterious. In our picture, different vacuum states are realized in different spacetime regions of the cosmological history.

1 Introduction

Except for providing explanations for neutrino-oscillation masses of neutrinos, dark matter, the baryon assymmetry and inflation, the standard model serves physicists extremely well for the moment. In our view the major motivation for seekng a theory beyond the standard model is the need for a theory that predicts the 20 or so free parameters of the standard model. Our by now old proposal of an assumption about the parameters, couplings and masses that can be appended to any proposed quantum field theory provides predictions for these free parameters. We call this assumption the Multiple Point Principle (MPP) [1,2,3]. The multiple point is a point in the action parameter space of a theory that is special in a way that is analogous to the way that the triple point of water is the special point in the phase diagram spanned by temperature and presure at which the solid, liquid and vapor phases of water coexist. The MPP states that fundamental physical parameters assume values that correspond to having a maximal number of different coexisting “phases” for the physically realized vacuum. There is phenomenological evidence suggesting that some or all of the about 20 parameters in

⋆ E-mail:bennett@alf.nbi.dk
⋆⋆ E-mail:hbech@alf.nbi.dk

the Standard Model (SM) that are not predicted within the framework of the SM correspond to the MPP values of these parameters. That these parameters take on special values (i.e., the multiple point values) poses from one viewpoint a fine tuning problem (why do constants of Nature take the MPP values). From another viewpoint, assuming the MPP as a law of Nature leads to a mechanism for fine-tuning. It is the latter viewpoint that is developed here. Moreover, we shall argue that a mild form of non locality is inherant to fine-tuning problems in general. We therefore develope a model for the relationship between fine-tuning, non-locality and the MPP.

2 Arguments for non-locality

2.1 From cosmological constant fine-tuning

Explaining the (dressed) value of the cosmological constant is an example of a fine tuning problem that would seem to require the breakdown of locality at least in a mild sense. As with any fine-tuning problem, the cosmological constant problem calls for a way to make the coupling dynamical in such a way that the values of such couplings are maintained at constant values (required for translational invariance). But this leads to a problem: if a coupling (e.g., the cosmological constant) is dynamical, the demands of a strictly local theory would be that the bare coupling can only depend on the spacetime point in question and indirectly on the past but certainly not on the future. However, if the bare cosmological constant (that is to be dynamically maintained at a constant value) immediately following the big bang is to already have its value fine-tuned once and for all - to say 120 decimal places - to the value that makes the dressed cosmological constant so small as suggested phenomenologically, we definitely have a problem with locality.

The problem is that the bare cosmological constant is relateable to the value of the dressed cosmological constant only if the details of the dressed cosmological constant (that did not exist when the bare value was already tuned to the valued required for the dressed vacuum) that will evolve in the future are known at the time of big bang[4]. We are forced to conclude that a strict principle of locality is not allowed if we want to have a dynamically maintained bare coupling and renormalization group corrections of a quantum field theory with a well-defined vacuum.

This suggests models with a mild form of non-locality consisting of an interaction that is the same between any pair of points in spacetime independent of the distance between these points. Assuming that this sort of non-locality is manifested through a non-local action $\hat{S}_{nl}$, this symmetry between any pair of space time points (i.e., identical interaction regardless of separation) is insured by requiring the invariance of $\hat{S}_{nl}$ under diffeomorphisms (reparameterization invariance). The non-local action $\hat{S}_{nl}$ is a function of functionals $I_{f_j}[\phi(x)]$: $\hat{S}_{nl} = \hat{S}_{nl}(\{I_{f_j}[\phi(x)\})$ where $I_{f_j}[\phi(x)] \stackrel{def}{=} \int dx^4 \sqrt{g(x)} f_j(\phi(x))$ and $f_j(\phi)$ might typically be a Lagrange density e.g., $f_j(\phi) = \mathcal{L}_j(\phi(x), \partial_\mu \phi(x))$. The symbol $\phi(x)$ stands for all the fields (and derivatives of same) of the theory.

An example of a nonlocal action would be any nonlinear function of the (reparameterization invariant) functionals $I_{f_i}, I_{f_j}, \cdots$; e.g., a term

$$\int d^4x \int d^4y \sqrt{g(x)g(y)}\phi^2(x)\phi^4(y).$$

Another example of a non-local (and nonlinear) action term more relevant to this paper is associated with having fixed values $I_{\mathtt{fixed}\ f_j}$ (fixed in the sense of being a law of Nature) of some extensive quantities $I_{f_j}[\phi]$. This amounts to having a δ-function term $\exp(S_{nl}(\{I_{f_j}\}) = \prod_j \delta(I_{f_j}[\phi] - I_{\mathtt{fixed}\ f_j})$ in the functional integration measure and results in the nonlocality that, strictly speaking, is inherant to any microcanonical ensemble (but which often is "approximated away" by using a canonical ensemble when phase space volume (or functional integration measure) is a sufficiently rapidly varying function of the extensive quantities).

An extensive quantity $I_{f_j}[\phi(x)]$ has a value for each imaginable Feynman path intgeral history of the Universe as it evolves from Big Bang to Big Crunch. The value $I_{\mathtt{fixed}\ f_j}$ is by assumption "frozen in" and cannot change during the lifetime of the Universe. This unchangeable "choice" $I_{\mathtt{fixed}\ f_j}$ then singles out a subset of all possible Feynman path integral histories that is consistent with the spacetime evoluion of our actually realized Universe having $I_{f_j}[\phi] = I_{\mathtt{fixed}\ f_j}$.

An interaction that is the same between the fields at *any* pair of spacetime points - regardless of separation - would not likely be perceived as a non-local interaction. Rather such spacetime omnipresent fields - a sort of background that is forever everywhere the same - would likely be interpreted as simply constants of Nature. This feature is reminiscent of baby universe theory the essence of which is that a physical constant can depend on something and still be a constant as a function of spacetime.

2.2 Arguments from short distance flucuation cancellation

Given that the prevailing feature of spacetime foam — the term coined by Wheeler to conjure up a picture of the Planck scale structure of spacetime — is a multitude of topologically nontrivial structures in a high curvature spacetime, it is natural to wonder how it comes about that spacetime at human distance scales of one meter say only deviate from being completely flat (i.e., ordinary Euclidean geometry) by tiny gravitational field effects that are almost negligible.

It is, however, not so trivial to see that short distance fluctuations will sum up to make large distance fluctuations in the curvature of spacetime. If we consider the parallel transport of a little vector from a genuine point in spacetime around a closed curve and back it is of course impossible that the fluctuation in angle of rotation caused by parallel transport along a long closed curve could be smaller than around a small closed curves of Planck size.

However, if we instead define an averaged effective geometry and consider a locally smeared way of parallel transporting and then parallel transport a vector defined in this smeared way along a long closed curve, it is no longer so that the fluctuation of the vector rotation after parallel transporting cannot be much smaller that the small (Planck) scale local fluctuations. However it is not immediately obvious what it means to define such an average of the geometry over

large volumes in order to average out small distance fluctuations. For instance it is not trivial to construct an approximately parallel vector field on a highly curved space.

3 The Multiple Point Principle (MPP)

The MPP was originally put forward in connection with theoretical predictions for the values of the three gauge coupling constants [1,2]. In addition to the assumption of the MPP, we also assumed in this first application of MPP our so-called Anti-GUT gauge group $G_{Anti-GUT}$ which consists of the 3-fold replication of the Standard Model Group (SMG): $G_{Anti-GUT} = SMG \otimes SMG \otimes SMG \stackrel{def}{=} SMG^3$ (in the extended version: $(SMG \times U(1))^3$) having one SMG factor for each generation of fermions and gauge bosons. We postulate that $G_{Anti-GUT}$ is broken to the diagonal subgroup (i.e., the usual SMG) at roughly the Planck scale.

In the original context of predicting the standard model gauge couplings using MPP (originally referred to as the principle of multiple point criticality), the principle asserts that the Planck scale values of the standard model gauge group couplings coincide with the multiple point, i.e., the point that lies in the boundary separating the maximum number of phases in the action parameter space corresponding to the gauge group $G_{Anti-GUT}$. The (Planck scale) predictions for the gauge couplings are subsequently will sum identified with the parameter values at the point in the action parameter space for the diagonal subgroup of $G_{Anti-GUT}$ that is inherited from the multiple point for $G_{Anti-GUT}$ after the Planck scale breakdown of the latter.

The idea was developed in the context of lattice gauge theory and the phases to which we refer are usually dismissed as lattice artifacts. (e.g., a Higgsed phase, a confined or Coulomb-like phase). Such phases have been studied extensively in the litterature for simple gauge groups and semi simple gauge groups with discrete subgroups (e.g. $SU(2)$ and $SU(3)$). One typically finds first order phase transitions between confined and Coulomb-like phases at critical values of the action parameters.

Taking such lattice artifact phases as physical reflects our suspicion that such phases are inherant to having a regulator. As a regulator in some form (be it a lattice, strings or whatever) is always needed for the consistency of any quantum field theory, it is consistent to assume the existence of a fundamental regulator. The "artifact" phases that arise in a theory with such a fundamental regulator (that we have chosen to implement as a fundamental lattice) are accordngly taken as ontological phases that have physical significance at the scale of the fundamental regulator (e.g., lattice). The assumption of an ontological fundamental regulator implies the existence of monopoles in terms of which the regulator induced phase can also be studied[5].

Finding the multiple point in an action parameter space corresponding to the gauge group $G_{Anti-GUT}$ is more complicated than for a single $SU(2)$ or $SU(3)$ say. The boundaries between phases in the action parameter space (i.e., the phase diagram) must be sought in a high dimensional parameter space essentially be-

cause $G_{Anti-GUT}$ being a non-simple group has many subgroups and invariant subgroups.

In fact there is a distinct phase for each subgroup pair (K, H) where K is a subgroup and H is an invariant subgroup such that $H \triangleleft K \subseteq G_{Anti-GUT}$. An element $U \in G_{Anti-GUT}$ can be parameterized as $U = U(g, k, h)$ where the Higgsed (gauge) degrees of freedom are elements g of the homogeneous space $G_{Anti-GUT}/K$. The (un-Higgsed) Coulomb-like and confined degrees of freedom are respectively the elements k of the factor group K/H and the elements $h \in H$.

4 A Familiar Analogy to the MPP as a Fine- Tuning Mechanism

Some important features of the MPP as a fine-tuning mechanism can be illustrated using an analogy to the familiar system in which the solid, liquid and vapour phases of water coexist. This occurs at the "triple point" of water, i.e., at the "triple point" values of temperature and presure. Because the transitions between these three phases are all first order, there is a whole range of combinations of the extensive variables energy and volume for which the system can only be realized by having the coexistence of the ice, liquid and vapour phases. But these three phases coexist only for the triple point values of temperature and presure, so there is a whole range of combinations of energy and volume that map onto the triple point values of the conjugate intensive variables temperature and presure with the result that these variables are fine-tuned to the triple point values. In this illustrative analogy, the triple point of water in the phase diagram spanned by the intensive parameters temperature and presure is analogous to the multiple point. As already stated, the multiple point is the (or a) point in the phase diagram that "touches" the maximun number of phases. In a phase diagram spanned by D intensive parameters (couplings), a generic multiple point can be in contact with up to $D + 1$ phases (in the illustrative example, $D = 2$ and the triple point is in contact with the $D + 1 = 3$ phases ice, liquid and vapour). In a non-generic situation, the multiple point can be in contact with more than $D + 1$ phases (e.g., accidently or due to symmetries).

For ease of illustration, consider now the even simpler system consisting of n_{H_2O} moles of H_2O in which just the ice and liquid phases coexist (at constant presure). Such a system is unavoidably realized (and the temperature fine-tuned to $0^o C = 273.15^o K$) for *any* value of the energy density $\rho_E = E/V_{n_{H_2O}}$ (E and $V_{n_{H_2O}}$ are respectively the energy and volume of the n_{H_2O} moles of H_2O) in the *finite* interval

$$\frac{n_{H_2O}}{V_{n_{H_2O}}} \int_{0^o\ K}^{273^o\ K} C_{p,ice}(T)dT < \rho_E < \frac{n_{H_2O}}{V_{n_{H_2O}}} \Bigg(\int_{0^o\ K}^{273^o\ K} C_{p,ice}(T)dT + (\text{molar heat of melting}) \Bigg) \quad (1)$$

($C_{p,ice}$ is the molar heat capacity of ice at constant pressure (e.g., 1 atm.)).

For any ρ_E in this interval, the system cannot be realized as a single phase but rather only as an equilibriated mixture of ice and liquid water. Even choosing ρ_E

at random there is a finite chance of landing in this interval in which case the temperature will be fine-tuned to 273.15° K.

5 The History of Our Universe as a Fine Tuner

Consider an analogy between the (3-dimensional) ice-water system with ρ_E in the interval of Eqn.(1) and our 4-dimensional universe with the value of an extensive variable $I_{f_j}[\phi(x)] \stackrel{def}{=} \int dx^4 \sqrt{g(x)} f_j(\phi(x))$ (with f_j any function of ϕ - see also Sec. 2 for notation) primordially fixed at a value $I_{fixed\ f_j}$ that can only be realized as a combination of two (for the sake of example - really there could be more than two) coexisting phases i.e., two degenerate vacuum states at field values that we denote as ϕ_{us} and ϕ_{other} where we take $\phi_{us} < \phi_{other}$. Here we are anticipating the introduction of an effective potential V_{eff} that has relative minima at the field values ϕ_{us} and ϕ_{other}. In 4-space, one generic possibility for having coexistent phases would be to have a phase with ϕ_{us} in an early epoch including say the universe as we know it and a phase with ϕ_{other} in a later epoch:

$$I_{fixed\ f_j} = f_j(\phi_{us})(t_{ignit} - t_{BB})V_3 + f_j(\phi_{other})(t_{BC} - t_{ignit})V_3 \tag{2}$$

where t_{ignit} is the "ignition" time (in the future) at which there is a first order phase transition from the vacuum at ϕ_{us} to the later vacuum at ϕ_{other}. V_3 is the 3-volume of the universe. The value of the "coupling constant" conjugate to $I_{fixed\ f_j}$ gets fine tuned (unavoidably by assumption of the coexistence of the two phases separated by a first order transition) by a mechanism that also depends on a phase that will first be realized in the future (at t_{ignit}). Such a mechanism is non-local. Note in particular that the right hand side of Eqn. 2 depends on t_{ignit}.

In order to formally define a "coupling constant" (intensive quantity) conjugate to some extensive quantity (e.g., $I_{fixed\ f_j}$), we introduce non-loality more abstractly. Let us restrict the non-local action $\hat{S}_{nl} = \hat{S}_{nl}(\{I_j[\phi(x)]\})$ to being a (also reparameterization invariant) non-local potential V_{nl} that is a function of (not necessarily independent) functionals

$$V_{nl} \stackrel{def}{=} V_{nl}(I_{f_i}[\phi], I_{f_j}[\phi], \cdots).$$

Define now an effective potential V_{eff} such that

$$\begin{aligned}\frac{\partial V_{eff}(\phi(x))}{\partial \phi(x)} &\stackrel{def}{=} \frac{\delta V_{nl}(\{I_{f_j}[\phi]\})}{\delta \phi(x)}|_{near\ min.} = \sum_i \left(\frac{\partial V_{nl}(\{I_{f_j}\})}{\partial I_{f_i}} \frac{\delta I_{f_i}[\phi]}{\delta \phi(x)} \right)|_{near\ min.} \\ &= \sum_i \frac{\partial V_{nl}(\{I_{f_j}\})}{\partial I_{f_i}}|_{near\ min.}\, f_i'(\phi(x))\end{aligned} \tag{3}$$

The subscript "near min" denotes the approximate ground state of the whole universe, up to deviations of $\phi(x)$ from its vacuum value (or vacuum values for a multi-phase vacuum) by any amount in relatively small spacetime regions. The solution to Eq. (3) is

$$V_{eff}(\phi) = \sum_i \frac{\partial V_{nl}(\{I_{f_j}\})}{\partial I_{f_i}} f_i(\phi) \tag{4}$$

We can identify the $\frac{\partial V_{nl}(\{I_{f_j}\})}{\partial I_{f_i}}$ as intensive quantities conjugate to the I_{f_i}.

Consider now the effective potential (4) in the special case that $V_{nl}(\{I_{f_j}\}) = V_{nl}(I_2, I_4) \stackrel{\mathrm{def}}{=} V_{nl}(\int d^4x\sqrt{g(x)}\phi^2(x), \int d^4y\sqrt{g(y)}\phi^4(y))$ in which case, (4) becomes

$$V_{eff} = \frac{\partial V_{nl}(I_2, I_4)}{\partial I_2}\phi^2(x) + \frac{\partial V_{nl}(I_2, I_4)}{\partial I_4}\phi^4(x) \stackrel{\mathrm{def}}{=} \frac{1}{2}m^2_{Higgs}\phi^2(x) + \frac{1}{4}\lambda\phi^4(x) \quad (5)$$

where the right hand side of this equation, which also defines the (intensive) couplings m^2_{Higgs} and λ, is recognised as a prototype scalar potential at the tree level. Of course the form of V_{nl} is, at least *a priori*, completely unknown to us, so - for example - the coupling constant m^2_{Higgs} cannot be calculated from Eqn. 5. The potential of Eqn. 5 with $m^2_{Higgs} < 0$ has an asymmetric minimun — at, say, the value ϕ_{us} resulting in spontaneous symmetry breakdown in the familiar way. This is just standard physics (without non-locality).

Actually we want to consider the potential V_{eff} having the two relative minima ϕ_{us} and ϕ_{other} - both at nonvanishing values of ϕ - alluded to at the beginning of this section. The second minimum comes about at a value $\phi_{other} > \phi_{us}$ when radiative corrections to (5) are taken into account and the top quark mass is not too large[6,7,3]. Which of these vacua - the one at ϕ_{us} or ϕ_{other} - would be the stable one in this two-minima Standard Model effective Higgs field potential depends on the value of m^2_{Higgs}. Since I_2 and I_4 are functions of t_{ignit} (as seen from Eqn. 2 with $f_j = \phi^2$ or ϕ^4), $m^2_{Higgs} \stackrel{\mathrm{def}}{=} \frac{\partial V_{nl}(\{I_2, I_4\})}{\partial I_2}$ is also a function of t_{ignit}.

Let us first use "normal physics" to see how the relative depths of the two minima of the double well are related to m^2_{Higgs} and to t_{ignit}. It can be deduced from[7] that a large negative value of m^2_{Higgs} corresponds to the relative minimum $V_{eff}(\phi_{other})$ being deeper than $V_{eff}(\phi_{us})$ (in which by assumption the Universe starts off following Big Bang) than for less negative values of m^2_{Higgs} (see Fig. 1). It can also be argued quite plausibly that a minimum in V_{eff} at ϕ_{other} much deeper than that at ϕ_{us} would correspond to an early (small) t_{ignit} inasmuch as the "false" vacuum at ϕ_{us} would be very unstable. However, as the value of the potential at ϕ_{other} approaches that at ϕ_{us}, t_{ignit} becomes longer and longer and approaches infinity as the values of V_{eff} at ϕ_{us} and ϕ_{other} become the same. The development of the double well potential and m^2_{Higgs} as a function of t_{ignit} is illustrated in Fig. 1. Note that the larger the difference $|V_{eff}(\phi_{other}) - V_{eff}(\phi_{us})|$ the more the realization of say $I_{fixed\ 2}$ will in general depend on t_{ignit}. If $V_{eff}(\phi_{us}) = V_{eff}(\phi_{other})$, t_{ignit} plays no role in realizing e.g. $I_{fixed\ 2}$ and the value of m^2_{Higgs} becomes independent of t_{ignit}.

6 Avoiding paradoxes arising from non-locality

In general the presence of non-locality leads to paradoxes. While the form that the non-local action (or potential V_{nl} in this discussion) is unknown to us, we make the 4 generically representative guesses portrayed as the 4 non-locality curves in Fig. 1. In particular, non-locality curves having a negative slope as a function of

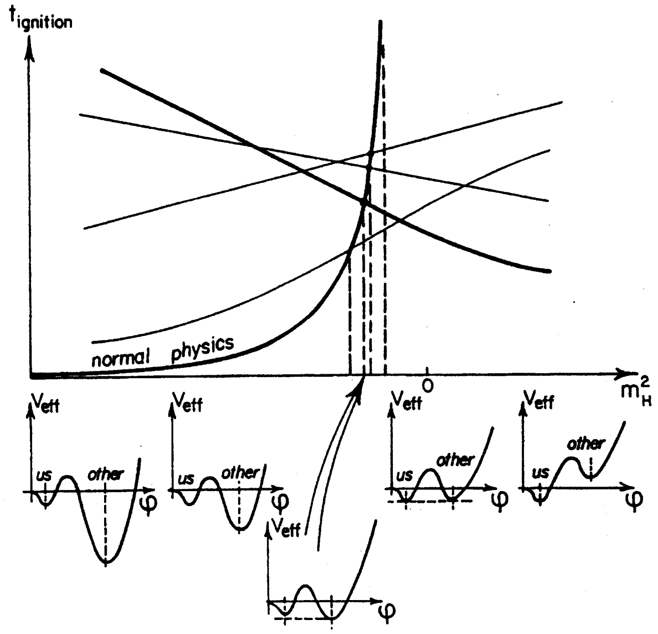

Fig. 1. The development of the double well potential and m_{Higgs} as a function of t_{ignit}. Note that all the more or less randomly drawn non-locality curves intersect the "normal physics" curve near where the vacua are degenerate (i.e., the MPP solution).

t_{ignit} lead to paradoxes in the following manner. Consider the non-locality curve in Fig. 1 drawn with bold line that is redrawn in a rotated position in Fig. 2. Let us make the assumption that t_{ignit} is large and see that this leads to a contradiction. Assuming that t_{ignit} is large, it is seen from the non-locality function in Fig. 2 (call it $m^2_{Higgs\ nl}(t_{ignit})$ to distinguish it from the "normal physics" $m^2_{Higgs}(t_{ignit})$) that this implies that the "normal physics" m^2_{Higgs} has a large negative value. But a large negative value of m^2_{Higgs} corresponds in "normal physics"to a (false) vacuum at ϕ_{us} that is very unstable and therefore to a very short t_{ignit} corresponding to a rapid decay to the stable vacuum at ϕ_{other}. So the paradox appears: the assumption of a *large* t_{ignit} implies a *small* t_{ignit}. This happens because in general $m^2_{Higgs}(t_{ignit}) \neq m^2_{Higgs\ nl}(t_{ignit})$ and is akin to the "matricide" paradox encountered for example when dealing with "time machines". It is well known[8,9,10] that Nature avoids such paradoxes by choosing a very clever solution in situations where these paradoxes lure.

In the case of the paradoxes that can come about due to non-locality of the type considered here, a clever solution that avoids paradoxes is available to Nature in the form of the Multiple Point Principle (MPP). The MPP solution corre-

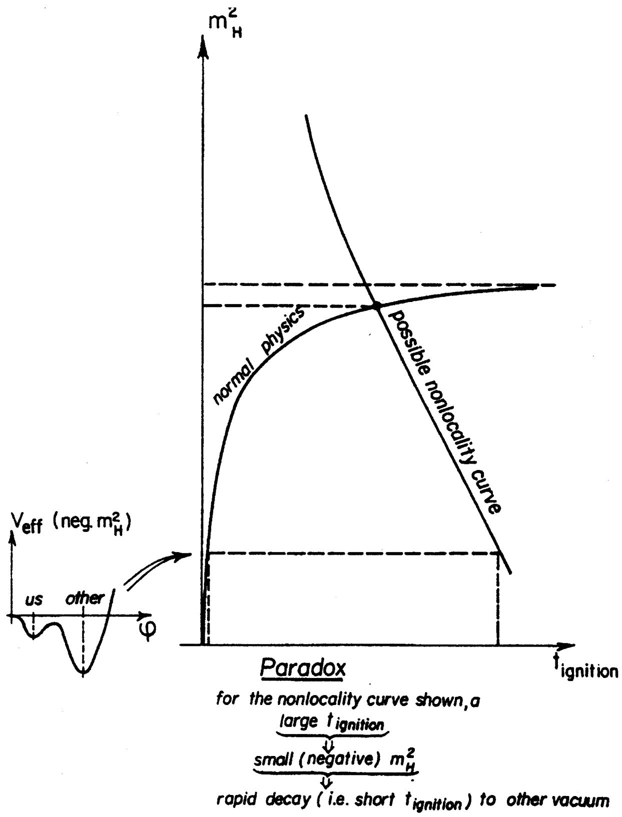

Fig. 2. Many non-locality curves could lead to paradoxes similar to the "matricide" paradox. Such paradoxes are avoided if the value of m_{Higgs} is fine-tuned to the multiple point critical value. This corresponds to the intersection of the "normal physics" curve with the "possible nonlocality" curve.

sponds to the intersection of the "normal physics" curve and the "non-locality curve" in Fig. 2. because here the vacua at ϕ_{us} and ϕ_{other} are (essentially) degenerate. But at this intersection point, $m^2_{Higgs}(t_{ignit}) = m^2_{Higgs\ nl}(t_{ignit})$ so the paradox is avoided. So the paradox is avoided at the multiple point. But at the multiple point, an intensive parameter has its value fine-tuned for a wide range of values of the conjugate extensive quantity. Fine-tuning can therefore be understood as a consequence of Nature's way of avoiding paradoxes that can come about due to non-locality.

It should be pointed out that the paradox-free solution corresponding to the intersection of the two curves in Fig. 2 occurs for a value of m^2_{Higgs} corresponding

to "our" vacuum at ϕ_{us} being very slightly unstable. The value of m^2_{Higgs} corresponding to the vacua at ϕ_{us} and ϕ_{other} being (precisely) degenerate is slightly less negative than that corresponding to the multiple point value of m^2_{Higgs} at the intersection of the curves. Note that the multiple point value of m^2_{Higgs} is very insensitive to which "guess" we use for the non-local action. Indeed all the "non-locality" curves in Fig. 1 intersect the "normal physics" curve at values of m^2_{Higgs} that are tightly nested together. The reason for this is that m^2_{Higgs} is a very slowly varying function of t_{ignit} as $m^2_{Higgs}(t_{ignit})$ approaches the value corresponding to degenerate vacua. The more nearly parallel the "normal physics" and the "non-locality" curves at the point of intersection, the less are the (paradoxical) effects of non-locality. For a point of intersection at values of t_{ignit} sufficiently large that (the "normal physics") $m^2_{Higgs}(t_{ignit}) \approx m^2_{Higgs}(\infty)$, the non-locality effects disappear as the curves become parallel since both curves become independent of t_{ignit}. If the curves were parallel, there would also be the possibility that these do not intersect in which case there would be no "miraculous solution" that could avoid the paradoxes imbued in having non-locality.

If the interval $|\phi^2_{other}$-$\phi^2_{us}|$ is large (e.g. of the order of the largest physically conceivable scale (Planck?) if tuning is to be maximally effective) and if $I_{fixed\ 2}$ falls not too close to the ends of this interval, then t_{ignit} will be something of the order of half the life of the universe. Actually, the approximate degeneracy of the vacua $V_{eff}(\phi_{us}) \approx V_{eff}(\phi_{other})$ may be characteristic of the temperature of the post-Big Bang universe in the present epoch and *not* characteristic of the high temperature that prevailed immediately following the Big Bang. At high temperatures, the free energy is considerably less than the total energy if the entropy is large enough. A phase with a large number of light particles - for example a Coulomb-like vacuum such as the "us" phase in which we live - could very plausibly be so strongly favoured at high temperatures that other phases - for example the "other" vacuum - simply disappeared at the high temperature of the universe immediately following the Big Bang.

If this were to have depleted the universe of the phase having ϕ_{other} at high temperatures, it would indeed be difficult to re-establish it in a lower temperature universe even if the vacuum at ϕ_{us} were to be only meta-stable and the vacuum at ϕ_{other} were the true vacuum at the lower temperature. Such an exchange of the true vacuum is indeed a possibility in going to lower temperatures inasmuch as the difference between the total energy and the free energy decreases in going to lower temperatures. Accordingly this difference becomes less effective in favouring a Coulomb-like phase at the expense of a phase with heavier particles.

At this point we point out that when we say the "vacuum at ϕ_{us}" and "vacuum at ϕ_{other}" we are thinking of the approximation $\phi = \phi_{us}$ and $\phi = \phi_{other}$ almost everywhere and at all times in respectively the early and late epochs of the universe in our discussion. More correctly we should talk about vacuum densities $\langle\phi(x^0, \mathbf{x})\rangle_{us}$ and $\langle\phi(x^0, \mathbf{x})\rangle_{other}$ where

$$\langle\phi(x^0, \mathbf{x})\rangle_{us} \stackrel{\text{def}}{=} \frac{1}{V_{us}} \int_{t_{BB}}^{t_{ignit}} dx^0 \int d^3\mathbf{x}\sqrt{g(x)}\phi(x^0, \mathbf{x})$$

where V_{us} denotes the the 4-volume of the universe in the first epoch. Density $\langle\phi(x^0, \mathbf{x})\rangle_{other}$ is defined analogously. The more correct $\langle\phi(x^0, \mathbf{x})\rangle_{us}$ is mentioned here so as not to confuse the reader when we talk about the changes in $V_{eff}(\phi_{us}(x^0, \mathbf{x}))$ as the universe cools following Big Bang.

Recall now that the value of say of $I_{fixed\ 2}$ can easily (i.e. as a generic possibility) assume a value that requires that the universe to be in the "phase" with $\langle\phi(x^0, \mathbf{x})\rangle_{other}$ during a sizeable part of its life if the universe is to have multiple point parameters in the course of its evolution (as required for avoiding the paradoxes that accompany non-locality). How can Nature overcome the energy barrier that must be surmounted in order to bring about the decay of the slightly unstable (false) vacuum with $\langle\phi(x^0, \mathbf{x})\rangle_{us}$ to the vacuum with $\langle\phi(x^0, \mathbf{x})\rangle_{other}$? Even producing just a tiny "seed" of the "true" vacuum having $\langle\phi(x^0, \mathbf{x})\rangle_{other}$ would be very difficult. What miraculously clever means can Nature devise so as to avoid deviations from a multiply critical evolution of the universe? One ingenious master plan that Nature may have implemented is the creation of life with the express "purpose" of evolving some (super intelligent?) physicists that could ignite a "vacuum bomb" by first creating in some very expensive accelerator the required "seed" of the "correct" vacuum having $\langle\phi(x^0, \mathbf{x})\rangle_{other}$ that subsequently would engulf the universe in a (for us) cataclysmic transition to the "other" phase thereby permitting the continued evolution of a "paradox-free" universe!

7 Conclusion

We attempt to justify the assertion that fine-tuning in Nature seems to imply a fundamental form of non-local interaction. This could be manifested in a phenomenologically acceptable form as everywhere in spacetime identical interactions between any pair of spacetime points. This would be implemented by requiring the non-local action to be diffeomorphism invariant.

Next we put forth our multiple point principle which states that coupling parameters in the Standard Model tend to assume values that correspond to the values of action parameters lying at the junction of a maximum number of regulator induced phases (e.g., so-called "lattice artifact phases") separated from one another in action parameter space by first order transitions. The action which of course is defined on a gauge group (e.g., the non-simple SM gauge group) governs fluctuation patterns along the various subgroup combinations (K, H) with $H \lhd K \subseteq G$ that characterize the phases that come together at the multiple point. We then consider extensive quantities that are functions of functionals $I_{f_j}[\phi(x)]$ that are essentially Feymann path histories of the Universe for functions $f_j(\phi)$ of the fields $\phi(x)$ and derivatives of these fields. We then think of the generic situation in which these extensive quantities can happen to be fixed at values that require the universe to be realized as two or more coexisting phases. We draw on the analogy to the forced coexistence of ice and liquid water that occurs for a whole range of possible total energies because of the finite heat of melting (first order phase transition). With our multiple point principle, the intensive quantities (couplings) conjugate to extensive quantities fixed in this way become fine-

tuned in a manner analogous to the fine tuning of temperature to 0^oC when the total energy of a system of H_2O can only be realized as coexisting ice and liquid phases.

One generic way of having coexisting phases in a quantum field theory in 3+1 dimensions would be to have different phases in different epochs of the lifetime of the Universe with phase transitions occuring at various times in the course of the lifetime of the Universe. If the transitions were first order, one would have fine-tuning of (intensive) couplings conjugate to extensive quantitity values that can only be realized by having coexisting (i.e., more than one) phases. But such a fine-tuning would involve non-locality: the fine-tuned values of coupling constants would depend on future phase transitions into phases that do not even exist at the time such couplings are fine-tuned.

Even non-locality of this sort (i.e., non-localy manifested as a diffeomorphism invariant non-local action) can lead to paradoxes of the "matricide paradox" type. We argue that such paradoxes are avoided when Nature chooses the multiple point principle solution to the problem of finetuning.

References

1. D. Bennett and H.B. Nielsen, Int. J. of Mod. Phys. **A9** (1994) 5155 .
2. D. Bennett and H.B. Nielsen, Int. J. of Mod. Phys. **A14** (1999) 3313.
3. C.D. Froggatt and H.B. Nielsen, Phys. Lett. **B368** (1996) 96.
4. N.C. Tsamis and R.P. Woodard, Ann. Phys.,**238** (1995) 1.
5. L.V. Laperashvili, H.B. Nielsen and D.A. Ryzhikh, Int. J. of Mod. Phys., **A18** (2003) 1.
6. S. Coleman and E. Weinberg, Phys. Rev. **D7** (1973) 1888.
7. M. Sher, Phys. Rep. **179** (1989) 273.
8. I.D. Novikov, Phys. Rev. **D45** (1992) 1989.
9. A. Lossev and I.D. Novikov, Class. Quantum Grav. **9** (1992) 2309.
10. C. Carlini, V.P. Frolov, M.B. Mensky, I.D. Novikov and H.H. Soleng, gr-qc/9506087.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 247)*
Portorož, Slovenia, July 12-17, 2003

Dynamics of Glue-Balls in N = 1 SYM Theory

Luzi Bergamin**

Technische Universität Wien
Institute of Theoretical Physics
Wiedner Hauptstrasse 8-10
1040 Wien, Austria

Abstract. The extension of the Veneziano-Yankielowicz effective Lagrangian with terms including covariant derivatives is discussed. This extension is important to understand glue-ball dynamics of the theory. Though the superpotential remains unchanged, the physical spectrum exhibits completely new properties.

1 Introduction

The low energy effective action of $N = 1$ SYM theory is written in terms of a chiral effective field $S = \varphi + \theta\psi + \theta^2 F$, which may be defined from the local source extension of the SYM action [1,2,3,4]

$$S \propto \frac{\delta}{\delta J} W[J, \bar{J}] \ , \qquad e^{iW[J,\bar{J}]} = \int \mathcal{D}V \ \ e^{i \int d^4x d^2\theta \ (J+\tau_0) \operatorname{Tr} W^\alpha W_\alpha + \text{h.c.}} \ . \tag{1}$$

With appropriate normalization S is equivalent to the anomaly multiplet $\bar{D}^{\dot{\alpha}} V_{\alpha\dot{\alpha}} = D_\alpha S$. $J(x)$ is the chiral source multiplet, with respect to which a Legendre transformation can be defined [3,4]. The resulting effective action is formulated in terms of the gluino condensate $\varphi \propto \operatorname{Tr} \lambda\lambda$, the glue-ball operators $F \propto \operatorname{Tr} F_{\mu\nu} F^{\mu\nu} + i \operatorname{Tr} F_{\mu\nu} \tilde{F}^{\mu\nu}$ and a spinor $\psi \propto (\sigma^{\mu\nu}\lambda)_\alpha F_{\mu\nu}$. An effective Lagrangian in terms of this effective field S has the form [1,2]

$$\mathcal{L}_{eff} = \int d^4\theta \ K(S, \bar{S}) - \left(\int d^2\theta \ S(\log \frac{S}{\Lambda^3} - 1) + \text{h. c.} \right) . \tag{2}$$

The correct anomaly structure is realized by the superpotential and thus $K(S, \bar{S})$ is invariant under all symmetries. In ref. [1] the explicit ansatz $K = k(\bar{S}S)^{1/3}$ had been made, which leads to chiral symmetry breaking due to $\langle S \rangle = \Lambda^3$, but supersymmetry is not broken as φ and ψ acquire the same mass $m = \Lambda/k$.

2 Glue-balls and constraint Kähler geometry

Though the spectrum found in ref. [1] does not include any glue-balls, such fields do appear in F. However, they drop out in the analysis of [1], as F is treated as an

** bergamin@tph.tuwien.ac.at

auxiliary field. Indeed, the highest component of a chiral superfield is auxiliary in standard SUSY non-linear σ-models, i.e. there appear no derivatives acting onto this field and moreover its potential is not bounded from below, but from above. In case of the Veneziano-Yankielowicz Lagrangian the part depending on the auxiliary field reads

$$\mathcal{L}_{aux} = k(\bar{\varphi}\varphi)^{-\frac{2}{3}}\bar{F}F + \left(\frac{1}{3}\varphi^{-\frac{2}{3}}\bar{\varphi}^{-\frac{5}{3}}F\bar{\psi}\bar{\psi} - F\log\frac{\varphi}{\Lambda^3} + \text{h. c.}\right) , \tag{3}$$

and the supersymmetric spectrum is obtained, *if and only if* F is eliminated by the algebraic equations of motion that follow from (3). This leads to the unsatisfactory result that glue-balls cannot be introduced in a straightforward way (cf. also [5]) which, in addition, contradicts available lattice-data [6].

However, in the special case of $N = 1$ SYM the elimination of F is not consistent: If F is eliminated from (3), this implies that the theory must be ultra-local in the field F *exactly*, i.e. even corrections to the effective Lagrangian which are not included in (2) are not allowed to change the non-dynamical character of F. If this field would be related to the fundamental auxiliary field, this restriction would be obvious. But in $N = 1$ SYM the situation is different: S is the effective field from a composite operator and F is not at all related to the fundamental auxiliary field D. As a consequence, the restriction of ultra-locality on F leads to an untenable constraint on the *physical* glue-ball operators (for details we refer to [4,7,8]).

As shown in ref. [2], the effective Lagrangian of [1] is not the most general expression compatible with all the symmetries, but the constant k may be generalized to a function $k(\frac{S^{1/3}}{\bar{D}^2\bar{S}^{1/2}}, \frac{\bar{S}^{1/3}}{D^2 S^{1/2}})$. This non-holomorphic part automatically produces space-time derivatives onto the field F, which is most easily seen when $K(S, \bar{S})$ is rewritten in terms of two chiral fields [8]:

$$K(S, \bar{S}) \to K(\Psi_0, \Psi_1; \bar{\Psi}_0, \bar{\Psi}_1) \tag{4}$$

Ψ_0 and Ψ_1 are not independent, but they must obey the constraints

$$\Psi_0 = S^{\frac{1}{3}} = \varphi^{\frac{1}{3}} + \frac{1}{3}\varphi^{-\frac{2}{3}}\theta\psi + \frac{1}{3}\theta^2(\varphi^{-\frac{2}{3}}F + \frac{1}{3}\varphi^{-\frac{5}{3}}\psi\psi) , \tag{5}$$

$$\Psi_1 = \bar{D}^2\bar{\Psi}_0 = \frac{1}{3}(\bar{\varphi}^{-\frac{2}{3}}\bar{F} + \frac{1}{3}\bar{\varphi}^{-\frac{5}{3}}\bar{\psi}\bar{\psi}) - \frac{i}{3}\theta\sigma^\mu\partial_\mu(\bar{\varphi}^{-\frac{2}{3}}\bar{\psi}) - \theta^2\Box\bar{\varphi}^{\frac{1}{3}} . \tag{6}$$

As F appears as lowest component of $\bar{\Psi}_1$, the Lagrangian includes a kinetic term for that field. In contrast to the situation in [1], this is not inconsistent as the potential in F may include arbitrary powers in that field (instead of a quadratic term only) and can be chosen to be bounded from below (instead of above). This way the field F is promoted to a usual physical field. It has been shown in [7] that there exist consistent models of this type. In [8] these ideas have been applied to $N = 1$ SYM, leading to an effective action of that theory with dynamical glue-balls as part of the low-energy spectrum. Formally, the effective potential looks the same as in the case of Veneziano and Yankielowicz:

$$\begin{aligned} V_{\text{eff}} = &-\tilde{g}_{\varphi\bar{\varphi}}F\bar{F} + \frac{1}{2}\tilde{g}_{\varphi\bar{\varphi},\bar{\varphi}}F(\bar{\psi}\bar{\psi}) + \frac{1}{2}\tilde{g}_{\varphi\bar{\varphi},\varphi}\bar{F}(\psi\psi) - \frac{1}{4}\tilde{g}_{\varphi\bar{\varphi},\varphi\bar{\varphi}}(\psi\psi)(\bar{\psi}\bar{\psi}) \\ &+ c\big(F\log\frac{\varphi}{\Lambda^3} + \bar{F}\log\frac{\bar{\varphi}}{\bar{\Lambda}^3} - \frac{1}{2\varphi}(\psi\psi) - \frac{1}{2\bar{\varphi}}(\bar{\psi}\bar{\psi})\big) \end{aligned} \tag{7}$$

However, in contrast to reference [1] the Kähler "metric"[1] is a function of φ *and* F, $\tilde{g}_{\varphi\bar{\varphi}}(\varphi, F; \bar{\varphi}, \bar{F})$. From eq. (7) the consistent vacua can be derived, for explicit expressions we refer to [8]. The most important properties of the Lagrangian (2) with (4) are:

The effective potential is minimized with respect to *all* fields φ, ψ and F. Consequently, the dominant contributions that stabilize the potential must stem from the Kähler part, not from the superpotential: The superpotential is a holomorphic function in its fields and therefore its scalar part must have unstable directions. In the present context there exists no mechanism to transform these instabilities into stable but non-holomorphic terms.

Though the model has the same superpotential as the Lagrangian of ref. [1] its spectrum is completely different: Chiral symmetry breaks by a vacuum expectation value (vev) of $\varphi \propto \Lambda^3$, but this mechanism is more complicated than in [1]. Any stable ground-state must have non-vanishing vev of F. But $\langle F \rangle$ is the order parameter of supersymmetry breaking and thus this symmetry is broken as well[2]. ψ is a massless spinor, the Goldstino.

The supersymmetry breaking scenario is of essentially non-perturbative nature[3]: it is not compatible with perturbative non-renormalization theorems, as the value of V_{eff} in its minimum and the vev of $T^{\mu}{}_{\mu}$ are no longer equivalent. In particular, the former can be negative, while the latter is positively semi-definite due to the underlying current-algebra relations. To our knowledge this is the first model, where this type of supersymmetry breaking has found a concrete description (cf. [7,8] for details).

Any ground state with $\langle \tilde{g}_{\varphi\bar{\varphi}} \rangle \neq 0$ can be equipped with stable dynamics for $p^2 < |\Lambda|^2$. In the construction of concrete kinetic terms it is important to realize that (4) may include expressions with explicit space-time derivatives. Again this is possible as F is not interpreted as an auxiliary field.

In summary, the Lagrangian of ref. [8] is the most general one, which can be formulated in terms of the effective field S. Consistent ground-states can be found together with broken supersymmetry only. It would be interesting to compare these results with a *different* action, which has supersymmetric ground-states. But the "pièce de résistance" for such an action is the fact, that it cannot start from the effective field S.

Acknowledgements

The author would like to thank U. Ellwanger, J.-P. Derendinger, E. Kraus, P. Minkowski, Ch. Rupp and E. Scheidegger for interesting discussions. This work has been supported by the Austrian Science Foundation (FWF) project P-16030-N08.

[1] This quantity is not equivalent to the true Kähler metric of the manifold spanned by Ψ_0 and Ψ_1, cf. [8].

[2] The author of ref. [2] concluded that this model cannot have a stable *supersymmetric* ground-state. This is in agreement with our results, as the model breaks down as $F \to 0$.

[3] The importance of such a breaking mechanism has been pointed out in [4] already, but a concrete description was not yet found therein.

References

1. G. Veneziano and S. Yankielowicz, Phys. Lett. **B113**, 231 (1982).
2. G. M. Shore, Nucl. Phys. **B222**, 446 (1983).
3. C. P. Burgess, J. P. Derendinger, F. Quevedo, and M. Quiros, Annals Phys. **250**, 193 (1996).
4. L. Bergamin, Eur. Phys. J. **C26**, 91 (2002).
5. G. R. Farrar, G. Gabadadze, and M. Schwetz, Phys. Rev. **D58**, 015009 (1998). D. G. Cerdeno, A. Knauf, and J. Louis, hep-th/0307198 (2003).
6. R. Peetz, F. Farchioni, C. Gebert, and G. Munster, hep-lat/0209065 (2002).
7. L. Bergamin and P. Minkowski, hep-th/0205240 (2002).
8. L. Bergamin and P. Minkowski, hep-th/0301155 (2003).

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 251)*
Portorož, Slovenia, July 12-17, 2003

Quantization of Systems with Continuous Symmetries on the Classical Background: Bogoliubov Group Variables Approach

Margarita V. Chichikina

Moscow State University Department of Quantum Field Theory and High Energy Physics Leninskie Gory 119992 Moscow Russian Federation
chich@goa.bog.msu.ru

1 Introduction

In present article the theory of quantization of nonlinear boson fields in a vicinity a nontrivial classical component by means of the Bogoliubov group variables is reprezented. Among canonical transformations of a boson field always there is an separation c-numerical component, therefore operator of a boson field always potentially contains a classical boson field. Therefore quantization of any nonlinear physical system should include quantization of a boson field in a vicinity of a classical background. If it is possible to suppose this field small, connected with it the effects, as a rule, are satisfactorily described in the framework of standart perturbation theory, when the classical field is as a first approximation considered equal to zero. If it is impossible to consider a classical field from the very beginning small (gravitation and extended particles), there is a task of the description of properties of physical system, in which main effect is the separation of a classical boson field.

The theories, existing on at present moment, of quantization on a classical background meet two basic difficulties. Firstly, — and this main — a problem of the conservation laws. The second difficulty of quantization on a classical background is the zero-mode problem. To bypass these problems we propose to use idea of the Bogoliubov, which was proposed in work of N.N.Bogoliubov. He has proposed to carry out quantization in the terms new variable. The transformations of the Bogoliubov are widely used in a quantum field theory. However, if the group of invariancy of system includes transformation of time, there are difficulties, connected with following: for reception of expression for Hamiltonian as generator of time translation it is necessary to know equations of motion, and for record of equations of motion in an explisit form it is necessary to know expression for Hamiltonian. This difficulty seriously limited a field of application of the Bogoliubov group variables, since did not enable to consider non-stationary systems.

In this work we propose quantizations, new a way, which will allow to use the Bogoliubov group variables for systems having arbitrary symmetry group

(provided that we we know representation of this group), including for non-stationary systems.

2 The basic ideas of a new method of quantization

In this section we represent schematically the basic ideas of quantization of boson fields by means of method of the Bogoliubov group variables. Firstly we pass to new variable - Bogoliubov group variables - and all considered quantities we express in the terms new variable.

Let variable $x^{'}$ are related to x by group transformation:

$$x' = F(x, a), \qquad x'' = F\left(F(x, a), b\right) = F(x, c), \qquad c = \varphi\left(a, b\right).$$

Variations of coordinates at a variation of parameters of group a are:

$$\left(\delta x'\right)^{i} = \xi^{i}_{s}(x')B^{s}_{p}(a)\delta a^{p},$$

where $i = 0, 1, 2, 3$- number of coordinate, $p = 1..., r$, and k - number of generators of group. The group properties of transformation are defined by tensor $B^{s}_{p}(a)$.

Let's define Bogoliubov transformation as follows:

$$f(x) = g\nu(x^{'}) + u(x^{'}),$$

dimensionless parameter g is supposed to be large, and $(u(x^{'}), a)$ are independent new variable (Bogoliubov group variable). Explisitly separated large component depends on $x^{'}$ as well as a quantum part. Thus we restore invariancy with respect to transformations group, which was broken by explisit separation of classical component, as it was pointed out in introduction. However consideration of variable τ as independent leads to the fact that right-hand side of our equation contains now on k variable more than left one. In order to equalize the number of variables we impose on $u(x^{'})$ some invariant conditions.

Secondly we develop the perturbation theory. All integrals of motions are expressed in new variables. Analyzing received expressions, we obtain conditions, at which application of perturbation theory is correct, namely: the classical part owes to satisfy to some equations. These equations turn out coinside with Eileur-Lagrange equations. However this result is not trivial: it was received not as a consequence of variational principle, but as a condition of development of a perturbative scheme for our system in the terms new variables, while for reception of these equations we did not need to know Hamiltonian structure as generator of transaltions on time.

So, at the second stage we obtain equations of motion for classical components.

And then we construct of system state space . At this stage of construction of our scheme we achieve correct number of degrees of freedom which will be equal to its real number. During realization of a reduction we obtain equations of motion for quantum correction for our field (the equations appear nonlinear). Also at this stage the explisit expressions for creation-annihilation operators of

quantums of a field are obtained as well as and spectrum of frequencies, that is actually, energy spectrum. Besides we obtain expressions for integrals of motion as generators of corresponding symmetry transformations:

$$O_0 = i\mathfrak{I}^{\alpha} \frac{\partial}{\partial \tau^{\alpha}},$$

In particular, hamiltonian is generator of shift on time. Let's note, that equations of motion of a classical part and explisit form of hamiltonian are received independently, that is is overcome basic difficulty of the description of non-stationary systems, which was described in introduction.

Also in zero order we shall obtain Heizenberg equation in the terms new variables.

Totality of the received results:

- Equation of motion of a classical part;
- Equation of motion of the quantum additive;
- Energy spectrum;
- Expressions for integrals of a motion;
- Heisenberg equations;
- Explicit expression for the field operator;

allows to assert, that our theory gives the complete description of system of boson fields, quantized on classical background. While theory guarantees exact performance of the conservation laws in any order of perturbation theory and also allows to avoid a zero-mode problem.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 254)*
Portorož, Slovenia, July 12-17, 2003

Singular Compactifications and Cosmology*

Laur Järv, Thomas Mohaupt and Frank Saueressig

Institute of Theoretical Physics, Friedrich-Schiller-University Jena,
Max-Wien-Platz 1, D-07743 Jena, Germany
`L.Jaerv, T.Mohaupt, F.Saueressig@tpi.uni-jena.de`

Abstract. We summarize our recent results of studying five-dimensional Kasner cosmologies in a time-dependent Calabi-Yau compactification of M-theory undergoing a topological flop transition. The dynamics of the additional states, which become massless at the transition point and give rise to a scalar potential, helps to stabilize the moduli and triggers short periods of accelerated cosmological expansion.

During the last year a lot of effort has been made to explain the astronomical evidence for an inflationary epoch of the early universe and the current modest accelerated expansion by invoking a scalar potential derived from string or M-theory compactifications. So far two mechanisms leading to potentials viable to describe accelerated cosmological expansion have been explored [1]: (i) compactifications on hyperbolic spaces [2] and (ii) compactifications with fluxes [3]. Our recent work [4,5] gives the first example of (iii) compactification on a singular internal manifold.

In the case of smooth compactifications one usually has a moduli space of vacua corresponding to the deformations of the internal manifold X and the background fields. For theories with eight or less supercharges this moduli space includes special points where X degenerates, rendering the corresponding low energy effective action (LEEA) discontinuous or singular. However, within the full string or M-theory these singularities are believed to be artifacts, which result from ignoring some relevant modes of the theory, namely the winding states of strings or branes around the cycles of X. Singularities of X arise when such cycles are contracted to zero volume, thereby introducing additional massless states. Incorporating these states leads to a smooth *gauged* supergravity action which entails a scalar potential.

In Calabi-Yau (CY) compactifications of M-theory undergoing a topological flop transition these additional states ('transition states') are given by N charged hypermultiplets which become massless at the transition locus. There are two ways to include the effect of these extra states in the LEEA. The usual LEEA is obtained by dimensional reduction on the smooth CY and contains only states which are generically massless. The flop manifests itself in a discontinuous change of the vector multiplet couplings at the transition locus. We call this descrip-

* Work supported by the 'Schwerpunktprogramm Stringtheorie' of the DFG.

tion the 'Out-picture' since the extra states are left out. On the contrary, the 'In-picture' is obtained by including the transition states as dynamical fields in the Lagrangian.

In [4] we constructed an In-picture LEEA for a generic M-theory flop by combining knowledge about the general $\mathcal{N} = 2, D = 5$ gauged supergravity action with information about the extra massless states.[1] While the vector multiplet sector could be treated exactly we used a toy model based on the quaternion-Kähler manifolds $\frac{U(1+N,2)}{U(1+N)\times U(2)}$ to describe the hypermultiplet sector. In order to find the gauging describing the flop we worked out the metrics, the Killing vectors, and the moment maps of these spaces. This data enabled us to construct a unique LEEA which has all the properties to model a flop: the extra hypermultiplets acquire a mass away from the transition locus while the universal hypermultiplet remains massless.

In [5] we considered an explicit model for a CY compactification undergoing a flop with $N = 1$ and investigated the effect of the transition states on five-dimensional Kasner cosmologies,[2]

$$ds^2 = -d\tau^2 + e^{2\alpha(\tau)}d\mathbf{x}^2 + e^{2\beta(\tau)}dy^2 \ . \tag{1}$$

Comparing the cosmological solutions of the Out- and the In-picture, we found that the inclusion of the dynamical transition states has drastic consequences for moduli stabilization and accelerated expansion.

As soon as we allow all light states to be excited the scalar fields no longer show the usual run-away behavior but are attracted to the flop region where they oscillate around the transition locus. Thus the "almost singular" manifolds close to the flop are dynamically preferred. This is somewhat surprising, because the potential has still many unlifted flat directions meaning there is no energy barrier which prevents the system from running away. Hence this effect cannot be predicted by just analyzing the critical points of the superpotential. The following thermodynamic analogy helps to explain the situation. Generically the available energy of the system is distributed equally among all the light modes ("thermalization"). Thus near the flop line the additional degrees of freedom get their natural share of it. Once this has happened, it becomes very unlikely that the system "finds" the flat directions and "escapes" from the flop region. Our numerical solutions confirm this picture: irrespective of the initial conditions the system finally settles down in a state where all the fields either approach finite values or oscillate around the transition region with comparable and small amplitudes. From time to time one sees "fluctuations from equilibrium", i.e., some mode picks up a bigger share of the energy for a while, but the system eventually thermalizes again. In an ideal scenario of moduli stabilization, however, one would like to have a damped system so that the moduli converge to fixed point values.

The second important aspect is that the scalar potential of the In-picture induces short periods of accelerated expansion in the three-space. Yet the net effect

[1] This strategy was first applied in [6] in the case of $SU(2)$ enhancement.

[2] This setup was previously considered in [7], but there the hypermultiplet manifold was taken to be hyper-Kähler which is not consistent with local supersymmetry.

of the accelerating periods on cosmic expansion is not very significant. Again, this feature can be understood qualitatively in terms of the properties of the scalar potential. The point is that the potential is only flat along the unlifted directions while along the non-flat ones it is too steep to support sustained accelerated expansion. Transient periods of acceleration occur when the scalar fields pass through their collective turning point, where running "uphill" the potential turns into running "downhill" and the potential energy momentarily dominates over the kinetic energy.[3] To get a considerable amount of inflation via a slow-roll mechanism, one would need to lift some of the flat directions gently without making them too steep.

In summary we see that the dynamics of the transition states is interesting and relevant, and can be part of the solution of the problems of moduli stabilization and inflation. One direction for further investigations is to consider more general gaugings of our five-dimensional model. Once gaugings leading to interesting cosmological solutions are found, one should clarify whether these can be derived from string or M-theory where they correspond to adding fluxes or branes. Another direction is to extend our construction to other topological transitions. In particular it would be interesting to study the effect of transition states on four-dimensional cosmologies arising, e.g., from type II compactifications on singular CY manifolds. It is conceivable that a realistic cosmology derived from sting or M-theory will have to include both the effects of fluxes and branes, and the possibility of internal manifolds becoming singular.

References

1. R. Emparan, J. Garriga: A note on accelerating cosmologies from compactifications and S branes, JHEP **0305** (2003) 026; P.K. Townsend: Cosmic acceleration and M-theory, hep-th/0308149.
2. P. Townsend, M.R.N. Wohlfarth: Accelerating cosmologies from compactification, Phys. Rev. Lett. **91** (2003) 061302; N. Ohta: Accelerating cosmologies from S-branes, Phys. Rev. Lett. **91** (2003) 061303; S. Roy: Accelerating cosmologies from M/string theory compactifications, Phys. Lett. **B567** (2003) 322; M.R.N. Wohlfarth: Accelerating cosmologies and a phase transition in M theory, Phys. Lett. **B563** (2003) 1; N. Ohta: A study of accelerating cosmologies from superstring / M theories, Prog. Theor. Phys. **110** (2003) 269; C.-M. Chen, P.-M. Ho, I.P. Neupane, J.E. Wang: A note on acceleration from product space compactification, JHEP **0307** (2003) 017; C.-M. Chen, P.-M. Ho, I.P. Neupane, N. Ohta, J.E. Wang: Hyperbolic space cosmologies, JHEP **0310** (2003) 058; M.R.N. Wohlfarth, Inflationary cosmologies from compactification?, hep-th/0307179; I.P. Neupane: Inflation from string/M-theory compactification?, hep-th/0309139.
3. L. Cornalba, M.S. Costa: A new cosmological scenario in string theory, Phys. Rev. **D66** (2002) 066001; A.R. Frey, A. Mazumdar: 3-form induced potentials, dilaton stabilization, and running moduli, Phys. Rev. **D67** (2003) 046006.
4. L. Järv, T. Mohaupt, F. Saueressig: Effective supergravity actions for flop transitions, hep-th/0310173.
5. L. Järv, T. Mohaupt, F. Saueressig: M-theory cosmologies from singular Calabi-Yau compactifications, hep-th/0310174.

[3] This behavior is also common to the models of hyperbolic and flux compactifications, where likewise the acceleration is not pronounced enough for primordial inflation.

6. T. Mohaupt, M. Zagermann: Gauged supergravity and singular Calabi-Yau manifolds, JHEP **0112** (2001) 026; J. Louis, T. Mohaupt, M. Zagermann: Effective actions near singularities, JHEP **0302** (2003) 053.
7. M. Brändle, A. Lukas: Flop Transitions in M-theory Cosmology, Phys. Rev. **D68** (2003) 024030.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard model *(p. 258)*
Portorož, Slovenia, July 12-17, 2003

Fundamental Physics and Lorentz Violation

Ralf Lehnert

CENTRA, Área Departamental de Física, Universidade do Algarve
8000-117 Faro, Portugal
rlehnert@ualg.pt

Abstract. The violation of Lorentz symmetry can arise in a variety of approaches to fundamental physics. For the description of the associated low-energy effects, a dynamical framework known as the Standard-Model Extension has been developed. This talk gives a brief review of the topic focusing on Lorentz violation through varying couplings.

1 Introduction

On the one hand, the Standard Model (SM) of particle physics is extremely successful phenomenologically. On the other hand, this conference is called *What Comes Beyond the Standard Model* because it is generally believed that the SM is really the low-energy limit of a more fundamental theory incorporating quantum gravity. Experimental research in this field faces various challenges. They include the expected Planck suppression of quantum-gravity signatures and the absence of a realistic underlying framework.

A promising approach for progress in quantum-gravity phenomenology is the identification of relations that satisfy three principle criteria: they must hold exactly in known physics, they are expected to be violated in candidate fundamental theories, and they must be testable with ultra-high precision. Spacetime symmetries satisfy all of these requirements. Lorentz and CPT invariance are key features of currently accepted fundamental physics laws, and they are amenable to Planck-sensitivity tests. Moreover, Lorentz and CPT breakdown has been suggested in a variety of approaches to fundamental physics. We mention low-energy emergent Lorentz symmetry [1], strings [2], spacetime foam [3], nontrivial spacetime topology [4], loop quantum gravity [5], noncommutative geometry [6], and varying couplings [7]. The latter of these mechanism will be discussed in more detail in this talk.

At presently attainable energies, Lorentz and CPT violating effects are described by a general extension of the SM. The idea is to include into the SM Lagrangian Lorentz and CPT breaking operators of unrestricted dimensionality only constrained by coordinate independence [8]. This Standard-Model Extension (SME) has provided the basis for many investigations placing bounds on Lorentz and CPT violation. For the best constraints in the matter and photon sectors, see Ref. [9] and Refs. [10,11], respectively. Note that certain Planck-suppressed SME operators for Lorentz and CPT breaking provide alternative ex-

planations for the baryon asymmetry in our universe [12] and the observed neutrino oscillations [13].

2 Lorentz violation through varying couplings

Early speculations in the subject of varying couplings go back to Dirac's numerology [14]. Subsequent theoretical investigations have shown that time-dependent couplings arise naturally in many candidate fundamental theories [15]. Recent observational claims of a varying fine-structure parameter α [16] have led to a renewed interest in the subject [17].

Varying couplings are associated with spacetime-symmetry violations. For instance, invariance under temporal and/or spatial translations is in general lost. Since translations are closely interwoven with the other spacetime transformations in the Poincaré group, one anticipates that Lorentz symmetry might be affected as well. This is best illustrated by an example. Consider the Lagrangian $\mathcal{L}$ of a complex scalar Φ, and suppose a spacetime-dependent parameter $\xi(x)$ is coupled to the kinetic term: $\mathcal{L} \supset \xi\partial_\mu\Phi\partial^\mu\Phi^*$. An integration by parts yields $\mathcal{L} \supset -\Phi(\partial_\mu\xi)\partial^\mu\Phi^*$. If, for instance, ξ varies smoothly on cosmological scales, $(\partial_\mu\xi) = k_\mu$ is essentially constant locally. The Lagrangian then contains a nondynamical fixed 4-vector k_μ selecting a preferred direction in the local inertial frame violating Lorentz symmetry.

The above example can be generalized to other situations. For instance, nonscalar fields can play a role, and Lorentz violation can arise through coefficients like k_μ in the equations of motion or in subsidiary conditions. Note that the Lorentz breaking is independent of the chosen reference frame: if $k_\mu \neq 0$ in a particular set of local inertial coordinates, k_μ is nontrivial in any coordinate system. In the next section, we show that varying couplings can arise through scalar fields acquiring expectations values in a cosmological context. Note, however, that the above argument for Lorentz violation is independent of the mechanism driving the variation of the coupling.

3 Four-dimensional supergravity cosmology

Consider a Lagrangian $\mathcal{L}$ with two real scalars A and B and a vector $F^{\mu\nu}$:

$$\frac{4\mathcal{L}}{\sqrt{g}} = \frac{\partial_\mu A\partial^\mu A + \partial_\mu B\partial^\mu B}{B^2} - 2R - MF_{\mu\nu}F^{\mu\nu} - NF_{\mu\nu}\tilde{F}^{\mu\nu} \ ,$$

$$M = \frac{B(A^2+B^2+1)}{(1+A^2+B^2)^2-4A^2} \ , \quad N = \frac{A(A^2+B^2-1)}{(1+A^2+B^2)^2-4A^2} \ , \tag{1}$$

where $g^{\mu\nu}$ represents the graviton and $g = -\det(g_{\mu\nu})$, as usual. We have denoted the Ricci scalar by R, the dual tensor is $\tilde{F}^{\mu\nu} = \varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}/2$, and the gravitational coupling has been set to one. Then, the Lagrangian (1) fits into the framework of the pure $N = 4$ supergravity in four spacetime dimensions.

To investigate Lagrangian (1) in a cosmological context, we assume a flat Friedmann-Robertson-Walker universe and model galaxies and other fermionic

matter by including the energy-momentum tensor $T^{\mu\nu}$ of dust, as usual.[1] In such a situation, the equations of motion can be integrated analytically [7] yielding a nontrivial dependence of A and B (and thus M and N) on the comoving time t. Comparison with the usual electrodynamics Lagrangian in the presence of a θ angle shows $\alpha \equiv 1/4\pi M$ and $\theta \equiv 4\pi^2 N$, so that the fine-structure parameter and the θ angle acquire related time dependences in our supergravity cosmology.

If mass-type terms $\mathcal{L}_m = -\sqrt{g}(m_A A^2 + m_B B^2)/2$ for the scalars are included into Lagrangian Eq. (1), our simple model can match the observed late-time acceleration of the cosmological expansion [7]. Note also that the scalars themselves obey Lorentz-violating dispersion relation [7,18].

Acknowledgments

I wish to thank C. Froggatt, N. Mankoč-Borštnik, and H.B. Nielsen for organizing a stimulating meeting. This work was supported in part by the the Centro Multidisciplinar de Astrofísica (CENTRA), the Fundação para a Ciência e a Tecnologia (FCT), and an ESF conference grant.

References

1. H.B. Nielsen and I. Picek, Nucl. Phys. B **211**, 269 (1983).
2. V.A. Kostelecký and S. Samuel, Phys. Rev. D **39**, 683 (1989); V.A. Kostelecký and R. Potting, Nucl. Phys. B **359**, 545 (1991); V.A. Kostelecký *et al.*, Phys. Rev. Lett. **84**, 4541 (2000).
3. G. Amelino-Camelia *et al.*, Nature (London) **393**, 763 (1998); D. Sudarsky *et al.*, Phys. Rev. D **68**, 024010 (2003).
4. F.R. Klinkhamer, Nucl. Phys. B **578**, 277 (2000).
5. J. Alfaro *et al.*, Phys. Rev. Lett. **84**, 2318 (2000).
6. S.M. Carroll *et al.*, Phys. Rev. Lett. **87**, 141601 (2001).
7. V.A. Kostelecký *et al.*, Phys. Rev. D, in press [astro-ph/0212003]; O. Bertolami *et al.*, astro-ph/0310344.
8. D. Colladay and V.A. Kostelecký, Phys. Rev. D **55**, 6760 (1997); Phys. Rev. D **58**, 116002 (1998); V.A. Kostelecký and R. Lehnert, Phys. Rev. D **63**, 065008 (2001).
9. D. Bear *et al.*, Phys. Rev. Lett. **85**, 5038 (2000).
10. S.M. Carroll, G.B. Field, and R. Jackiw, Phys. Rev. D **41**, 1231 (1990).
11. V.A. Kostelecký and M. Mewes, Phys. Rev. Lett. **87**, 251304 (2001); Phys. Rev. D **66**, 056005 (2002).
12. O. Bertolami *et al.*, Phys. Lett. B **395**, 178 (1997).
13. V.A. Kostelecký and M. Mewes, hep-ph/0308300; hep-ph/0309025.
14. P.A.M. Dirac, Nature (London) **139**, 323 (1937).
15. See, e.g., E. Cremmer and J. Scherk, Nucl. Phys. B **118**, 61 (1977); P. Forgács and Z. Horváth, Gen. Rel. Grav. **10**, 931 (1979); A. Chodos and S. Detweiler, Phys. Rev. D **21**, 2167 (1980); J.D. Barrow, Phys. Rev. D **35**, 1805 (1987).
16. J.K. Webb *et al.*, Phys. Rev. Lett. **87**, 091301 (2001).
17. See, e.g., J.-P. Uzan, Rev. Mod. Phys. **75**, 403 (2003).
18. R. Lehnert, Phys. Rev. D **68**, 085003 (2003).

[1] The dust can be accommodated into the supergravity framework, which also contains fermions uncoupled from the scalars.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 261)*
Portorož, Slovenia, July 12-17, 2003

Functional Approach to Squeezed States in Non-commutative Theories

Lubo Musongela

The Abdus Salam International Centre for Theoretical Physics
High Energy section
strada costiera , 11
P.O.Box 586
34100 Trieste
Italy

1 Summary

We show how the difference of structure between the classical Poisson brackets and the quantum commutators for the non commutative plane generically leads to a harmonic oscillator whose position mean values are not strictly periodic. We also show that no state saturates simultaneously all the non trivial Heisenberg uncertainties in this context.

This raises the question of the nature of quasi classical states in this model. We propose an extension based on a variational principle.

2 Periodicity of the harmonic oscillator.

The non commutative plane attracted interest when it was realized it could appear in the context of string theory [1]. It is defined by the commutation relations

$$[\hat{x}_1, \hat{x}_2] = i\,\theta \quad , \quad [\hat{x}_j, \hat{p}_k] = i\,\hbar\,\delta_{jk} \quad ; \quad \theta, \hbar \quad > 0 \quad .$$

One can realize this algebra in a simple way:

$$\hat{x}_1 = i\hbar\partial_{p_1} - \frac{1}{2}\frac{\theta}{\hbar}p_2 \quad , \quad \hat{x}_2 = i\hbar\partial_{p_2} + \frac{1}{2}\frac{\theta}{\hbar}p_1 \quad ,$$
$$\hat{p}_1 = p_1 \quad , \quad \hat{p}_2 = p_2 \quad , \quad \langle\phi|\psi\rangle = \int d^2p\,\phi^*(p)\,\psi(p) \quad .$$

The time evolution of any operator in quantum mechanics is governed by the equation $\dot{\hat{A}} = i\hbar^{-1}[\hat{H}, \hat{A}]$. Solving it for the position operator x_1 and computing its mean value in an arbitrary state, one finds [2]

$$\langle\hat{x}_1(t)\rangle = \frac{1}{\hbar(\lambda_1 - \lambda_2)(\lambda_1 + \lambda_2)}\;(c_1\cos\lambda_1 t + s_1\sin\lambda_1 t + c_2\cos\lambda_2 t + s_2\sin\lambda_2 t)$$

where the frequencies are given by

$$\lambda_{1,2} = \pm\sqrt{\frac{k}{m} + \frac{k^2\theta^2}{2\hbar^2} \pm \frac{k^{3/2}\theta}{2\hbar^2\sqrt{m}}\sqrt{4\hbar^2 + km\theta^2}} \quad .$$

The coefficients c_i, s_i depend on the state. The mean value of the position x_1 will not be periodic unless the ratio of the two frequencies is rational. This is at odds with the usual quantum mechanic result and originates from the fact that here the commutator $[\hat{x}_1, \hat{x}_2] = i\theta$ has not exactly the form of the Poisson bracket $\{\hat{x}_1, \hat{x}_2\} = 0$.

3 Conflicting equalities

Another difference of usual Q.M and non commutative theories lies in the fact that in the later ones, all the non trivial uncertainties can not be satisfied simultaneously [3]. Habitually, the non trivial commutators imply only canonically conjugate pairs like x_1, p_1. The squeezed or coherent states saturate the corresponding bounds: $\Delta x_i \Delta p_i = \hbar/2$. On the non commutative plane, a state saturating all the non trivial uncertainties would obey the equations

$$\Delta x_1 \Delta x_2 = \frac{\theta}{2} \Longrightarrow \hat{a}_1|\psi\rangle = (\hat{x}_1 + i\lambda_1\hat{x}_2 + \mu_1)|\psi\rangle = 0 \quad ,$$
$$\Delta x_1 \Delta p_1 = \frac{\hbar}{2} \Longrightarrow \hat{a}_2|\psi\rangle = 0 \quad , \quad \Delta x_2 \Delta p_2 = \frac{\hbar}{2} \Longrightarrow \hat{a}_3|\psi\rangle = 0 \quad ,$$

where $\hat{a}_2$ and $\hat{a}_2$ are similar to $\hat{a}_1$. It is easily obtained that $[\hat{a}_2, \hat{a}_3]|\psi\rangle = i\theta|\psi\rangle = 0$ so that no state saturates simultaneously all the three bounds. This is a second difference with usual quantum mechanics.

4 Generalization based on a functional approach.

In usual quantum mechanics, the squeezed states saturate the uncertainty relations and so realize a minimum of the functional $\sum_k (\Delta x_k \Delta p_k - h/2)^2$. We have verified that our proposal works in the usual theory: we recover the known gaussian functions and, besides them, other states which can be expressed as products of gaussians with specific hypergeometrics [2]. The generalization we consider here also rely on a functional involving only the non trivial commutators:

$$S = \hbar^2\left(\Delta x_1 \Delta x_2 - \frac{1}{2}\theta\right)^2 + \theta^2\left(\Delta x_1 \Delta p_1 - \frac{1}{2}\hbar\right)^2 + \theta^2\left(\Delta x_2 \Delta p_2 - \frac{1}{2}\hbar\right)^2$$
$$+ \lambda\left(\langle\psi|\psi\rangle - 1\right) \quad .$$

The normalization of the state is achieved thanks to a Lagrange multiplier.

Varying the action, one obtains that the desired wave function is an eigenfunction of the operator

$$\mathcal{O} = \bar{a}_1\partial_{p_1}^2 + \bar{a}_2\partial_{p_2}^2 + i\left(\bar{a}_1\frac{\theta}{\hbar^2}p_2 + a_3\right)\partial_{p_1} + i\left(-\frac{\theta}{\hbar^2}\bar{a}_2 p_1 + a_4\right)\partial_{p_2}$$
$$+ \left(\bar{a}_5 p_1^2 + \bar{a}_6 p_2^2 + a_7 p_1 + a_8 p_2 + a_9\right) \quad .$$

with zero eigenvalue. The $\bar{a}_i$ are constants. They are related to the physical characteristics of the state. Some exact solutions can be found. For example, introducing the variables y_1, y_2 by the relations

$$p_1 = a_1 y_1 - \frac{a_4}{a_2^2} \quad , \quad p_2 = a_2 y_2 + \frac{a_3}{a_1^2} \quad ,$$

the operator takes the simpler form

$$\mathcal{O} = -\partial_{y_1}^2 - \partial_{y_2}^2 + it_1^2(-y_2\partial_{y_1} + y_1\partial_{y_2}) + (t_2^2 y_1^2 + t_3^2 y_2^2 + t_4 y_1 + t_5 y_2 + t_6) \quad .$$

One in particular find Gaussian solutions

$$\psi(p) = N \exp\left(c_1 y_1^2 + c_2 y_2^2 + c_3 y_1 y_2 + c_4 y_1 + c_5 y_2\right)$$

provided that the relation $2c_1 + 2c_2 + c_4^2 + c_5^2 - t_6 = 0$ holds. One can similarly find other solution under other assumptions [2].

It should be stressed that our use of a functional to define meaningful states is similar to the one of [4], although the physical requirements in the two cases are different.

5 Conclusions

We have suggested an approach towards squeezed states which relies on a functional method involving non trivial Heisenberg uncertainty relations. It is a generalization of the squeezed states of usual quantum mechanics. We have found special solutions to the second order differential equations obtained on the non commutative plane.

One of the crucial points which remain to be addressed is the nature of the critical points found here. To know if these states are maxima or minima of the action, one has to ressort to a second order analysis. However, as the most general solution of the second order partial differential equations involved are not known, such a computation cannot tell us by itself if we are in front of an absolute minimum.

Another crucial question about the states found by the method presented here is the other properties of the usual coherent states they may possess, like completeness [5,6,7,8,9]. If this was the case, they might be legitimate candidates for the definition of a physically meaningful star product [10]. The fact that some solutions obtained are special functions which are solutions of Sturm-Liouville systems is promising.

References

1. N.Seiberg, E.Witten, JHEP 0002 (2000) 020
2. Musongela Lubo, hep-th/0304039.
3. K.Bolonek, P.Kosinski, Phys.Lett. **B547** (2002) 51-54
4. S.Detournay, Cl.Gabriel, Ph.Spindel, Phys.Rev. **D66** (2002) 125004
5. A. Perelomov, *Generalized coherent states and their application*, Springer (1986).

6. V.I.Manko, G.Marmo, E.C.G Sudarchan, F.Zaccaria, Physica Scripta **55** (1997) 528.
7. G.Vitiello, Phys. Rev. Lett. **66** (1991) 2056
8. C.A. Nelson and M.Fields, Phys. Rev **A51** (1995) 2410
9. G. Alexanian, A. Pinzul, A. Stern, Nucl.Phys. **B600** (2001) 531-547
10. A.Voros, Phys. Rev. **A40** (1989) 6814

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 265)*
Portorož, Slovenia, July 12-17, 2003

Constraining the Curvaton Scenario

Marieke Postma

The Abdus Salam ICTP, Strada Costiera 11, 34100 Trieste, Italy
postma@ictp.trieste.it

Abstract. We analyse the curvaton scenario in the context of supersymmetry. Supersymmetric theories contain many scalars, and therefore many curvaton candidates. To obtain a scale invariant perturbation spectrum, the curvaton mass should be small during inflation $m \ll H$. This can be achieved by invoking symmetries, which suppress the soft masses and non-renormalizable terms in the potential. Other model-independent constraints on the curvaton model come from nucleosynthesis, gravitino overproduction, and thermal damping. The curvaton can work for masses $m \gtrsim 10^4$ GeV, and very small couplings (e.g. $h \lesssim 10^{-6}$ for $m \lesssim 10^8$ GeV).

1 The curvaton scenario

It is now widely believed that the early universe went through a period of rapid expansion, called inflation. In addition to explaining the homogeneity and isotropy of the observable universe, inflation can provide the seeds for structure formation. This makes models of inflation predictive, but also restrictive. The observed, nearly scale-invariant perturbation spectrum requires very small coupling constants and/or masses, which renders many models unnatural. For this reason it is worthwhile to explore alternative ways of producing density perturbations.

In the curvaton scenario, the adiabatic perturbations are not generated by the inflaton field, but instead result from isocurvature perturbations of some other field — the *curvaton* field. After inflation the isocurvature perturbations have to be converted into adiabatic ones. Such a conversion takes place with the growth of the curvaton energy density compared to the total energy density in the universe. This alternative method of producing adiabatic did not attract much attention until recently [1].

The usual implementation of the curvaton scenario is the following. If the curvaton is light with respect to the Hubble constant during inflation, it will fluctuate freely, leading to condensate formation. In the post-inflationary epoch the field remains effectively frozen until the Hubble constant becomes of the order of the curvaton mass, $H \sim m_\phi$, at which point the curvaton starts oscillating in the potential well. During oscillations, the curvaton acts as non-relativistic matter, and its energy density red shifts slower than the radiation bath. Hence, the ratio of curvaton energy density to radiation energy density grows $\rho_\phi/\rho_\gamma \propto a$, with a the scale factor of the universe, and isocurvature perturbations are transformed into curvature perturbations. This conversion halts when the curvaton comes to dominate the energy density, or if this never happens, when it decays.

It seems natural to try to embed the curvaton scenario within supersymmetric (SUSY) theories. SUSY theories contain many flat directions, and therefore many possible curvaton candidates. The problem, however, is that during inflation SUSY is broken dynamically and soft mass terms are generated, which are typically of the order of the Hubble constant. But this is no good: $m_\phi \sim H$ during inflation leads to a large scale dependence of the produced perturbations, in conflict with observations. A way out of this is to invoke symmetries. Soft mass terms are for example suppressed in D-term inflation, and in no-scale supergravities. Another possibility is to identify the curvaton with a pseudo-Goldstone bosons.

2 Constraints

There are several model independent constraints on the curvaton scenario. We will discuss them briefly here; see the original paper for more details [3].

First of all, the curvaton scenario should give rise to the observed spectrum of density perturbations. Curvature perturbations $\mathcal{R}$ of the correct magnitude are obtained for $\mathcal{R} \approx (f/3\pi)(H_*/\phi_*) \approx 5 \times 10^{-5}$ [1]. Here the subscript $*$ denotes the quantity at the time observable scales leave the horizon. Further, $f = \rho_\phi/\rho_{\rm tot}$ evaluated at the time of curvaton decay. If the curvaton contributes less than 1% to the total energy density, i.e., $f < 0.01$, then the perturbations have an unacceptable large non-Gaussianity. If during inflation $m_\phi \ll H$ — which is required to get a nearly scale invariant perturbation spectrum — quantum fluctuations of the curvaton grow until $m_\phi^2 \langle \phi^2 \rangle \sim H^4$, with an exponentially large coherence length. We will assume that this sets the initial curvaton amplitude $\phi_* \sim \sqrt{\langle \phi^2 \rangle}$. The non-detection of tensor perturbations puts an upper bound on the Hubble scale during inflation $H_* \lesssim 10^{14}$ GeV. Finally, in the curvaton scenario the adiabatic density perturbations can be accompanied by isocurvature perturbations in the densities of the various components of the cosmic fluid. There are particularly strong bounds on the isocurvature perturbations in cold dark matter.

The initial curvaton amplitude should be $\phi_0 \lesssim M_P$, to avoid a period of curvaton driven inflation. Stronger constraints are obtained if non-renormalizable operators are taken into account: $V_{\rm NR} = |\lambda|^2 {M_P}^{-n} \phi^{4+n}$. Non-renormalizable operators are unimportant for small enough masses, $m_\phi \lesssim m_{\rm eff} = {V_{\rm eff}}''$. For larger masses, the curvaton slow-rolls in the non-renormalizable potential during and after inflation. In the post-inflationary epoch this leads to a huge damping of the fluctuations, making it is impossible to obtain the observed density contrast within the context of the curvaton scenario [4].

The curvaton scenario should not alter the succesful predictions of big bang nucleosynthesis (BBN). This implies that the curvaton should decay before the temperature drops below MeV. To avoid gravitino overproduction requires a reheat temperature $T_R \lesssim 10^9$ GeV. This also constrains the curvaton scenario, since isocurvature perturbations are converted in adiabatic perturbations only after inflaton decay.

Finally, one should take into account various thermal effects. Large thermal masses, $m_{\rm th} \gtrsim m_\phi$, induce early oscillations, which are generically fatal for

the curvaton scenario. In addition, thermal evaporation of the curvaton scenario should be avoided.

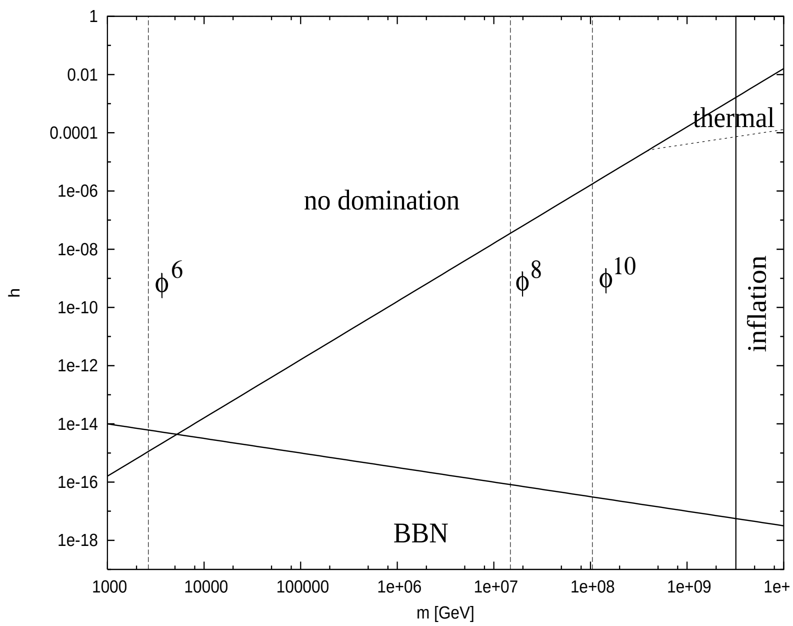

Fig. 1. Parameter space for curvaton domination ($f \gtrsim 0.5$). In this plot T_R is unconstrained. The constraints from BBN, domination, non-renormalizable terms, ϕ-dominated inflation, and thermal damping are shown.

3 Results

Figs. 1 and 2 shows the parameter space for *curvaton domination*: $f \gtrsim 0.5$. In the figure 1 the reheating temperature is arbitrary high, whereas in the figure 2 the gravitino constraint is taken into account. In all parameter space $m_\phi \sim 10^{-4} H_*$. Models with $V_{NR} \sim \phi^{4+n}/M_P{}^n$ and $n \leq 2$ are ruled out. For higher values of n, the curvaton scenario can be succesfull for curvaton masses in the range 10^4 GeV $\lesssim m_\phi \lesssim 10^9$ GeV. Couplings have to be small $h \lesssim 10^{-6}$, even $h \lesssim 10^{-10}$ if the gravitino constraint is taken into account. Thermal effects can be neglected for such small couplings.

One can ask whether there are any natural canditates for the curvaton. Moduli and other fields with only Planck suppressed couplings generically decay after big bang nucleosynthesis, thereby spoiling its succesfull predicitions. MSSM flat directions typically have too small masses and too large couplings to play the rôle of the curvaton. Better curvaton candidates are the right-handed sneutrino and the Peccei-Quinn axion, which can have large masses and small couplings. In all cases though, considerable tuning of parameters is needed.

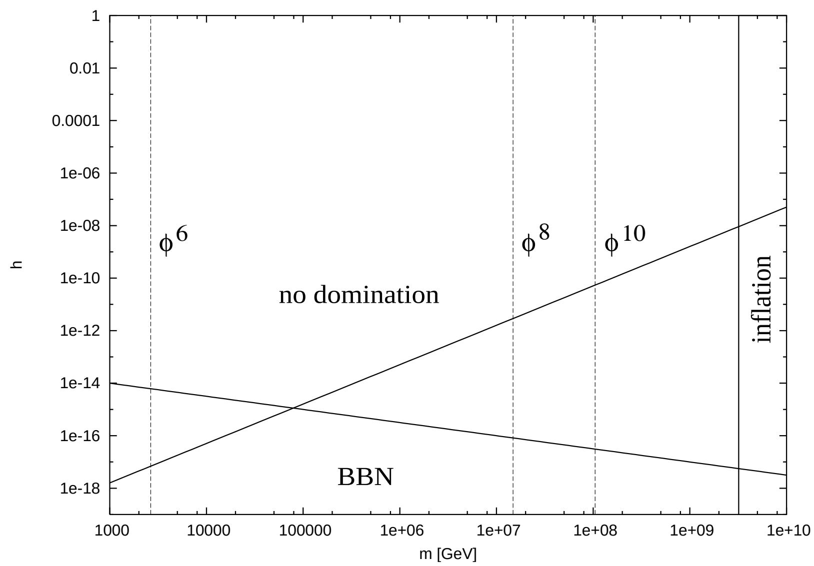

Fig. 2. Parameter space for curvaton domination ($f \gtrsim 0.5$). In this plot $T_R \lesssim 10^9$ GeV. The constraints from BBN, domination, non-renormalizable terms, ϕ-dominated inflation, and thermal damping are shown.

Acknowledgments

This work was supported by the European Union under the RTN contract HPRN-CT-2000-00152 Supersymmetry in the Early Universe.

References

1. K. Enqvist and M. S. Sloth, Nucl. Phys. B **626**, 395 (2002) [arXiv:hep-ph/0109214]. D. H. Lyth and D. Wands, Phys. Lett. B **524**, 5 (2002), [arXiv:hep-ph/0110002]; T. Moroi and T. Takahashi, Phys. Lett. B **522**, 215 (2001)
2. D. H. Lyth, C. Ungarelli and D. Wands, arXiv:astro-ph/0208055.
3. M. Postma, Phys. Rev. D **67**, 063518 (2003) [arXiv:hep-ph/0212005].
4. K. Enqvist, A. Jokinen, S. Kasuya and A. Mazumdar, arXiv:hep-ph/0303165.

BLED WORKSHOPS
IN PHYSICS
VOL. 4, NO. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 269)*
Portorož, Slovenia, July 12-17, 2003

D-Branes and Unitarity of Noncommutative Field Theories

Alessandro Torrielli

Dipartimento di Fisica "Galileo Galilei"
Università di Padova
I.N.F.N. Sezione di Padova
torriell@pd.infn.it

Abstract. We review the original result we have obtained in the analysis of the breaking of perturbative unitarity in space-time noncommutative field theories in the light of their relations to D-branes in electric backgrounds

Since the seminal work of Seiberg and Witten[1] it was realized that open strings in the presence of an antisymmetric constant background are effectively described at low energy by certain noncommutative field theories, identified by a precise set of Feynman rules[2]. The derivation by Seiberg and Witten was originally made for magnetic backgrounds only, corresponding to space-space noncommutativity. For electric backgrounds it is well known that problems are present when the electric field approaches a critical value E_{cr}, beyond which the string develops a classical instability[3,4,5]: tachyonic masses appear in the spectrum both for neutral (which is the case we will consider here) and charged open strings, connected with the vanishing of the effective tension and an uncontrolled growth of the oscillation amplitude of the modes in the direction parallel to the field. For purely charged strings this phenomenon (which has no analog in particle field theory) coexists with the quantum instability due to pair production which is the analog of the Schwinger phenomenon in particle electrodynamics[1]. When one tries to perform the Seiberg-Witten zero-slope limit to noncommutative field theories one reaches a point in which the ratio between the electric field and its critical value becomes greater than one, and the string enters the region of instability. We will show that precisely at this point a flip-mechanism produces the appearance of the tachyonic branch cut due to the closed string sector in the non-planar diagrams, which is responsible for the lack of unitarity of the limit amplitude[2]. We relate the breaking of perturbative unitarity of these space-time

[1] As a note, we recall that for neutral strings ($+q$-charge on one end, $-q$ on the other), one has $E_{cr} = 1/(2\pi\alpha'|q|)$. For charges $q_1 \neq q_2$ on the two boundaries, one finds that the pair production rate diverges at the same critical value of the classical instability $E_{cr} = 1/(2\pi\alpha'|max q_i|)$

[2] Space-time noncommutative field theories present besides a series of exotic behaviours[6,7]

noncommutative theories with the fact that they coincide with those obtained by conveying string theory in an unstable vacuum. This vacuum is likely to decay in the full string theory to a suitable configuration of branes.

We take as action for a bosonic open string attached to a D-brane lying in the first $p+1$ dimensions with an antisymmetric constant background on its world-volume the following

$$S = \frac{1}{4\pi\alpha'} \int_{C_2} d^2 z \,(g_{ij} \partial_a X^i \partial^a X^j - 2i\pi\alpha' B_{ij} \epsilon^{ab} \partial_a X^i \partial_b X^j). \tag{1}$$

At one loop the string world-sheet is the cylinder $C_2 = \{0 \leq \Re w \leq 1, w \equiv w + 2i\tau\}$.

If one sets $w = x+iy$, the relevant propagator on the boundary of the cilinder $(x = 0, 1)$ can be written as[8]

$$G(y, y') = \frac{1}{2}\alpha' g^{-1} \log q - 2\alpha' G^{-1} \log \Big[\frac{q^{\frac{1}{4}}}{D(\tau)} \vartheta_4 (\frac{|y-y'|}{2\tau}, \frac{i}{\tau})\Big], \quad x \neq x', \tag{2}$$

$$G(y, y') = \frac{\pm i\theta}{2} \epsilon_\perp (y - y') - 2\alpha' G^{-1} \log \Big[\frac{1}{D(\tau)} \vartheta_1 (\frac{|y-y'|}{2\tau}, \frac{i}{\tau})\Big], \quad x = x', \tag{3}$$

where $q = e^{-\frac{\pi}{\tau}}$, $\pm$ correspond to $x = 1$ and $x = 0$ respectively, and $\epsilon_\perp(y) = sign(y) - \frac{y}{\tau}$. The open string parameters are as in [1]. With this propagator and the suitable modular measure, the non-planar two-point function can be written as follows:

$$A_{1.2} = \mathcal{N} {G_s}^2 \, 2^{\frac{3d}{2}} \pi^{\frac{3d}{2}-2} {\alpha'}^{\frac{d}{2}-3} \int_0^\infty dt\, t^{1-\frac{d}{2}} \Big[\eta(\frac{it}{2\pi\alpha'})\Big]^{2-d} \times$$
$$e^{-\frac{\pi^2 {\alpha'}^2}{t} k g^{-1} k} \int_0^1 d\nu \left[\frac{e^{-\frac{\pi^2\alpha'}{2t}} \vartheta_4 (\nu, \frac{2\pi i \alpha'}{t})}{\frac{2\pi\alpha'}{t} [\eta(\frac{2\pi i\alpha'}{t})]^3}\right]^{-2\alpha' k G^{-1} k} . \tag{4}$$

Here $\mathcal{N}$ is the normalization constant, $d = p+1$, G_s is the open string coupling constant. We have omitted the traces of the Chan-Paton matrices.

This is a form suitable for the field theory limit. We perform the zero-slope limit of Seiberg and Witten, which consists in sending $\alpha' \to 0$ keeping θ and G fixed (in this case we will keep t and ν fixed as well). This can be done setting $\alpha' \sim \epsilon^{\frac{1}{2}}$ and the closed string metric $g \sim \epsilon$, and then sending $\epsilon \to 0$ [1]. One obtains [8]

$$A^{lim}_{1.2} = \mathcal{N} 2^{\frac{3d}{2}} \pi^{\frac{3d}{2}-2} {g_f}^2 \int_0^\infty dt t^{1-\frac{d}{2}} e^{-m^2 t + k\theta G \theta k/4t} \int_0^1 d\nu e^{-t\,\nu(1-\nu)\, k G^{-1} k}. \tag{5}$$

This reproduces the expression for the two-point function in the noncommutative ϕ^3 theory[9,10,11] with coupling constant $g_f = G_s \, {\alpha'}^{\frac{d-6}{4}}$. We choose now $d = 2$. This case is peculiar, since the tachyon mass $m^2 = \frac{2-d}{24\alpha'}$ goes to zero. In two dimensions the background is an electric one and noncommutativity is necessarily space-time.

In order to interpret the field theory analysis, one continues the expression (4) in the complex variable $k^2 = kG^{-1}k$. It has a branch cut driven by the small-t (closed-channel) behavior

$$\exp\Big(-k^2\,\frac{\pi^2\alpha'^2\,[1-(E/E_{cr})^2]}{t}\Big). \tag{6}$$

The field-theory limit retains in turn another branch cut $Re[S] = Re[-k^2] > 0$ from the large-t behaviour

$$\exp\Big(-k^2\,t\,\nu(1-\nu)\Big). \tag{7}$$

$E_{cr} = \frac{g}{2\pi\alpha'}$ is the critical value of the electric field[3]. The small-t cut is on the physical side $Re[S] = Re[-k^2] > 0$. But we see that it is driven by a quantity which changes sign if the electric field overcomes its critical value, that is when the string enters the classical instability region. This is exactly what the Seiberg-Witten limit produces, since it scales $(E/E_{cr})^2 \sim (\alpha'/g)^2 \sim 1/\epsilon$. In the field theory limit the branch cut flips to the unphysical side[12], and in fact the amplitude (5) for $m^2 = 0$ has two branch cuts, one of which is physical, the other one tachyonic, coalescing at the origin. By treating m^2 as a free parameter independent of d and by taking it positive[4], one can see the shift of (6) to a cut starting from $S > 4m^2$ which gives the physical pair-production cut, and a closed branch cut starting at $S > 4m^2/[1-(E/E_{cr})^2]$, which flips in the limit.

Acknowledgments

I wish to thank of heart Antonio Bassetto and Roberto Valandro. I wish to thank also L. Bonora and R. Russo for useful discussions, and C.-S. Chu, J. Gomis, S. Kar and M. Tonin for suggestions. Finally I wish to thank the organizers of the Euresco conference in Portoroz for such a beatiful and inspiring meeting.

References

1. N. Seiberg and E. Witten, *JHEP* **9909**, 032 (1999) [hep-th/9908142].
2. T. Filk, *Phys. Lett. B* **376**, 53 (1996).
3. C. P. Burgess, *Nucl. Phys. B*, **294**, 427 (1987).
4. C. Bachas and M. Porrati, *Phys. Lett. B* **296**, 77 (1992) [hep-th/9209032].
5. N. Seiberg, L. Susskind and N. Toumbas, *JHEP* **0006**, 021 (2000) [hep-th/0005040].
6. N. Seiberg, L. Susskind and N. Toumbas, *JHEP* **0006**, 044 (2000) [hep-th/0005015].
7. J. L. Barbon and E. Rabinovici, *Phys. Lett. B* **486**, 202 (2000) [arXiv:hep-th/0005073].
8. O. Andreev and H. Dorn, *Nucl. Phys. B* **583**, 145 (2000) [hep-th/0003113].
9. M. R. Douglas and N. A. Nekrasov, *Rev. Mod. Phys.* **73**, 977 (2001) [hep-th/0106048].
10. R. J. Szabo, *Phys. Rept.* **378**, 207 (2003) [hep-th/0109162].
11. A. Torrielli, PhD thesis, Padua University, 2003 [hep-th/0301091].
12. A. Torrielli, Phys. Rev. D **67**, 086010 (2003) [hep-th/0207148].

[3] We have taken unitary charges (see (1)) and a closed string metric $g_{\mu\nu} = g\,\eta_{\mu\nu}$

[4] This is the point of view of many works by Di Vecchia et al. in deriving field theory amplitudes from the bosonic string, and it is also the approach usually adopted in the noncommutative case

Bled Workshops
in Physics
Vol. 4, No. 2–3

Proceedings to the Euroconference on
Symmetries beyond the Standard
model *(p. 272)*
Portorož, Slovenia, July 12-17, 2003

Spinorial Cohomology and Supersymmetry

Dimitrios Tsimpis

Department of Mathematics
King's College
The Strand
London WC2R 2LS
United Kingdom

Abstract. We review recently developed cohomological methods relating to the study of supersymmetric theories and their deformations.

String theory reduces at low energies to ordinary field theory supplemented by an infinite tower of higher-order curvature (derivative) corrections organised in a double-series expansion in α' and $g_{\mathtt{string}}$. These corrections can in principle be derived by string-theoretical perturbative methods, but in certain cases supersymmetry alone is restrictive enough to detrmine them.

The question is more sharply posed in the case of M-theory (see [1] for a review) whose low-energy effective field-theory limit is captured by ordinary eleven-dimensional supergravity [2]. In order to gain insight into the nature of M-theory, one needs to go beyond this limiting approximation; for example, by including supersymmetric higher-order curvature corrections. In the absence of an underlying perturbative analogue of string theory such corrections cannot be found systematically, even in principle, but one might hope that supersymmetry would be sufficient to determine at least the first-order correction.

Higher-order corrections have far-reaching implications to a variety of physical problems at the heart of our understanding of M-theory: they can be used to derive modifications to macroscopic black hole properties such as the entropy-area formula and to test duality conjectures, like the AdS/CFT correspondence, beyond the leading-order approximation; they may give a mechanism for moduli stabilisation and they may provide a way to bypass no-go theorems concerning dS vacua. However, even the first such deformation –corresponding to canonical mass dimension six operators (R^4)– has proven notoriously difficult to determine.

The construction of the complete R^4 terms in type II string theory or in M-theory remains an open problem, see [3] for a review and further references. A recent attempt was made in [4] to construct an R^4 action in type IIB based on a chiral measure in on-shell IIB superspace, but it was subsequently observed that such a measure does not exist [5]. A completely different approach was taken by the authors of [3] who used partial results from type II string theory and attempted to lift them to eleven dimensions, but this proved to be too difficult to carry out completely.

Recently-developed techniques relating to the study of supersymmetric theories, which go under the name of spinorial cohomology, offer perhaps the most promising and comprehensive approach to tackling these long-standing issues. Spinorial cohomology, which is the subject of this brief review, was introduced in [6] and was further elaborated in [7]. It provides a powerful and systematic way of organising the field content and the possible deformations of supersymmetric theories, as well as explicitly determining higher-order curvature corrections. Pure-spinor cohomology [8], which was recently used by Berkovits in a covariant approach to string theory [9,10], has been shown to be isomorphic to the concept of spinorial-cohomology with unrestricted coefficients [11,12].

Spinorial-cohomology has already found applications in a variety of contexts; in studies of ten-dimensional maximally-supersymmetric Yang-Mills theories [6,13,14], eleven-dimensional supergravity [7,12,15], and the world-volume effective theory of the M2 brane [16]. In [13,16] the first-order curvature correction to $D = 10$, $N = 1$ SYM and to the M2-brane action respectively, was given explicitly. In [12] a method was outlined which makes it possible to read off the first deformation of eleven-dimensional supergravity from the five-brane anomaly-cancelling term. Moreover, it was argued that the supersymmetric completion of this term is the unique anomaly-cancelling invariant at this dimension which is at least quartic in the fields.

In order to give a geometrical definition of spinorial cohomology (see [12] for a more detailed discussion), we will suppose that we have the usual machinery of supergravity, i.e. Lorentzian structure group, connection, torsion and curvature. In particular we suppose that the tangent bundle is a direct sum of the odd and even bundles. The space of forms admits a natural bigrading according to the degrees of the forms and their Grassmann character. The space of forms with p even and q odd components is denoted by $\Omega^{p,q}$. The exterior derivative d maps $\Omega^{p,q}$ to $\Omega^{p+1,q} + \Omega^{p,q+1} + \Omega^{p-1,q+2} + \Omega^{p+2,q-1}$. Following [17] we split d into its various components with respect to the bigrading

$$d = d_0 + d_1 + \tau_0 + \tau_1 \ , \tag{1}$$

where $d_0(d_1)$ is the even (odd) derivative with bidegrees $(1, 0)$ and $(0, 1)$ respectively, while τ_0 and τ_1 have bidegrees $(-1, 2)$ and $(2, -1)$. These two latter operators are purely algebraic and involve the dimension-zero and dimension-three-halves components of the torsion tensor respectively. The fact that $d^2 = 0$ implies in particular that $\tau_0^2 = 0$. We can therefore consider the cohomology of τ_0 and set

$$H_\tau^{p,q} = \{\omega \in \Omega^{p,q} | \tau_0\omega = 0 \bmod \omega = \tau_0\lambda, \lambda \in \Omega^{p+1,q-2}\} \, . \tag{2}$$

We can now define a spinorial derivative d_F which will act on elements of $H_\tau^{p,q}$. If $\omega \in [\omega] \in H_\tau^{p,q}$ we set

$$d_F[\omega] := [d_1\omega] \, . \tag{3}$$

It is easy to check that this is well-defined, i.e. $d_1\omega$ is τ_0-closed, and $d_F[\omega]$ is independent of the choice of representative. Furthermore it is simple to check that $d_F^2 = 0$. This means that we can define the spinorial cohomology groups

$$H_F^{p,q} := H^{p,q}(d_F|H_\tau) \ , \tag{4}$$

in the obvious fashion.

To illustrate the meaning of the cohomology groups let us now turn to $D = 10$, $N = 1$ SYM. The two lowest-dimensional components of the Bianchi identity are

$$d_1 F_{0,2} + \tau_0 F_{1,1} = 0 \qquad (5)$$

$$d_1 F_{1,1} + \tau_0 F_{2,0} = 0 \, . \qquad (6)$$

The ordinary (undeformed) theory is obtained by setting $[F_{0,2}] = 0$ and this is equivalent to specifying an element of $H_F^{0,1}$. Hence we can set $F_{0,2} = 0$ in which case $F_{1,1}$ determines an element of $H_F^{1,1}$. One can easily see that this cohomology group is a spinor superfield λ^α satisfying the constraint $D_\alpha \lambda^\beta = (\gamma^{ab})_\alpha{}^\beta F_{ab}$, i.e. it is the on-shell field strength supermultiplet.

If we relax the ordinary constrain $F_{\alpha\beta} = 0$ by introducing a current $J_{\alpha\beta}$ on the right-hand side, then the latter must be spinorially closed by virtue of the Bianchi identity. On the other hand, if it is trivial, the connection can be redefined to regain the original equations of motion. This implies that $H_F^{0,2}$ describes the currents of the theory (which are in one-to-one correspondence with the antifields). If we are interested in looking at deformations of the theory, we need to consider the same cohomology group but with the coefficients restricted to be tensorial functions of the field strength superfield λ and its derivatives. We denote this group by $H_F^{0,2}(phys)$. The two groups are quite different; the former is dual to the physical fields while the latter describes a composite multiplet in the theory. The group $H_F^{0,2}(phys)$ has been explicitly computed in [6] for the nonabelian SYM at order α'^2 to give the most general supersymmetric action at this order [13]. It has also been computed for the abelian SYM at order α'^3 to prove rigorously the absence of any supersymmetric deformation [14].

Eleven-dimensional supergravity can also be understood in terms of spinorial cohomolgy groups [7,12]. Apart from a purely geometrical formulation in terms of the supertorsion [18], the theory admits a formulation in terms of a closed four-form in superspace. From this point of view, physical fields are elements of $H_F^{0,3}$, while deformations are parametrized by the $H_F^{0,4}(phys)$ spinorial-cohomology group [7,12].

One can also give a description in terms of a 'dual' super four-form obeying a deformed Bianchi identity [19,20]

$$dG_7 = \frac{G^2}{2}_4 + \beta X_8 \, , \qquad (7)$$

where β is a parameter of dimension ℓ^6 and $X_8 = \mathrm{Tr}(R^4) - \frac{1}{4}(\mathrm{Tr}(R^2))^2$. This equation uses imput from M-theory, namely the five-brane anomaly cancellation condition, and it can be used to read off the deformation consistent with X_8. One can then make contact with the four-form formulation by solving the BI's up to dimension -1 [12].

The solution of the supertorsion Bianchi identities of eleven-dimensional supergravity will be presented in [15], completing the work initiated in [20]. Moreover, provided a nontrivial element of the spinorial-cohomology group $H_F^{0,4}(phys)$

has been determined through the procedure described above, the information can be fed in the results of [12,15] to give the explicit expression of the deformed equations of motion –and therefore the first-order (R^4) correction to the Lagrangian– of eleven-dimensional supergravity. This procedure, albeit tedious, is straightforward and can be carried out in practice.

Acknowledgements

I would like to acknowledge the support of EU contract HPRN-2000-00122 and PPARC grants PPA/G/S/1998/00613 and PPA/G/O/2000/00451.

References

1. M. J. Duff, "The World In Eleven Dimensions: Supergravity, Supermembranes And M-Theory," Bristol, UK: IOP (1999).
2. E. Cremmer, B. Julia and J. Scherk, "Supergravity Theory In 11 Dimensions," Phys. Lett. B **76** (1978) 409.
3. K. Peeters, P. Vanhove and A. Westerberg, "Supersymmetric higher-derivative actions in ten and eleven dimensions, the associated superalgebras and their formulation in superspace," hep-th/0010167, Class. Quant. Grav. **18** (2001) 843.
4. S. de Haro, A. Sinkovics and K. Skenderis, "A supersymmetric completion of the R^4 term in IIB supergravity," hep-th/0210080.
5. N. Berkovits and P.S. Howe, unpublished.
6. M. Cederwall, B. E. Nilsson and D. Tsimpis, "The structure of maximally supersymmetric Yang-Mills theory: Constraining higher-order corrections," hep-th/0102009, JHEP **0106** (2001) 034.
7. M. Cederwall, B. E. Nilsson and D. Tsimpis, "Spinorial cohomology and maximally supersymmetric theories," hep-th/0110069, JHEP **0202** (2002) 009.
8. P. S. Howe, *Pure spinors, function superspaces and supergravity theories in ten-dimensions and eleven-dimensions*, Phys. Lett. B **273** (1991) 90.
9. N. Berkovits, *Super-Poincare covariant quantization of the superstring*, hep-th/0001035, JHEP **0004** (2000) 018.
10. N. Berkovits, "Covariant quantization of the supermembrane," hep-th/0201151, JHEP **0209** (2002) 051.
11. P. S. Howe, unpublished.
12. P. S. Howe and D. Tsimpis, "On higher-order corrections in M theory," hep-th/0305129, JHEP **0309** (2003) 038.
13. M. Cederwall, B. E. Nilsson and D. Tsimpis, "D = 10 super-Yang-Mills at $O(\alpha'^2)$," hep-th/0104236, JHEP **0107** (2001) 042.
14. M. Cederwall, B. E. Nilsson and D. Tsimpis, "Spinorial cohomology of abelian d = 10 super-Yang-Mills at $O(\alpha'^3)$," hep-th/0205165, JHEP **0211** (2002) 023.
15. M. Cederwall, U. Gran, B. E. Nilsson and D. Tsimpis, to appear.
16. P. S. Howe, S. F. Kerstan, U. Lindstrom and D. Tsimpis, "The deformed M2-brane," hep-th/0307072, JHEP **0309** (2003) 013.
17. L. Bonora, K. Lechner, M. Bregola, P. Pasti and M. Tonin, "A Discussion Of The Constraints In N=1 Sugra-Sym In 10-D," Int. J. Mod. Phys. A **5** (1990) 461.
18. P. S. Howe, "Weyl superspace," hep-th/9707184, Phys. Lett. B **415** (1997) 149.
19. M. J. Duff, J. T. Liu and R. Minasian, "Eleven-dimensional origin of string / string duality: A one-loop test," hep-th/9506126, Nucl. Phys. B**452** (1995) 261.

20. M. Cederwall, U. Gran, M. Nielsen and B. E. Nilsson, "Manifestly supersymmetric M-theory," hep-th/0007035, JHEP **0010** (2000) 041; "Generalised 11-dimensional supergravity," hep-th/0010042.

BLEJSKE DELAVNICE IZ FIZIKE, LETNIK 4, ŠT. 2–3, ISSN 1580–4992

BLED WORKSHOPS IN PHYSICS, VOL. 4, NO. 2–3

Zbornik konference 'Euroconference on Symmetries Beyond the Standard Model', Portorož, 12. – 17. julij 2003

Proceedings to the 'Euroconference on Symmetries Beyond the Standard Model', Portorož, July 12. – 17., 2003

Uredili Norma Mankoč Borštnik, Holger Bech Nielsen, Colin D. Froggatt in Dragan Lukman

Publikacijo sofinancira Ministrstvo za šolstvo, znanost in šport

Tehnična urednica Andreja Čas

Založilo: DMFA – založništvo, Jadranska 19, 1000 Ljubljana, Slovenija

Natisnila Tiskarna MIGRAF v nakladi 100 izvodov

Publikacija DMFA številka 1551